

BLOSUM62						
Gapop 10.0 , Gapext 0.5						
BLOSUM62						
Query	Score	Start	End	AlnLen	AlnQual	AlnQual2
1 MAGIPGLLFLFLCAVGQ.....LKYAQICYWIKGNYLDREG	383	1	107	384	100	100
US-09-072-384-18	2080	1	107	384	100	100
Sequence:						
Protein - protein search, using sw model						
on on:	December 2, 2001, 16:24:08	;	Search time 32.13 Seconds			
			(without alignments)			
	437.037	Million	cell updates/sec			
table:						
Effect score:						
oring table:						
searched:	100059	seqs,	36664827	residues		
Total number of hits satisfying chosen parameters:		100059				
Minimum DB seq length:	0					
Maximum DB seq length:	2000000000					
st-processing:	Minimum Match 0%					
	Maximum Match 100%					
	Listing first 45 summaries					
Database :	SwissProt_39;*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
Result No.	Score	Query	Match	Length	DB ID	Description
1	194	9.3	316	1	GSEP_BACLI	P80057 bacillus 11
2	135.5	6.5	313	1	MPP_BACSU	P3790 bacillus su
3	124	6.0	266	1	ELI_BOVIN	Q2153 bos taurus
4	121.9	5.8	490	1	FALO_RABBT	O10045 oryctolagus
5	119	5.7	269	1	EL2_PIG	P08419 sus scrofa
6	115.5	5.6	269	1	EL2A_HUMAN	P0217 homo sapien
7	115	5.6	271	1	EL2_MOUSE	P0208 mus musculus
8	115	5.5	376	1	FALO_TROCA	P81428 tropidectis
9	114.5	5.5	238	1	TRV5_AEDEA	P2787 aedes aegypti
10	113	5.4	266	1	ELI_PIG	P00772 sus scrofa
11	113	5.4	266	1	ELI_RAT	P0773 ratus norvegicus
12	112.5	5.4	259	1	TRV5_STRE	P0775 streptomyces
13	112.5	5.4	273	1	YDGD_ECOLI	P76176 escherichia
14	111.5	5.4	498	1	FALO_HUMAN	P0742 homo sapien
15	111.5	5.4	875	1	NEPR_HUMAN	P56730 homo sapien
16	111	5.3	253	1	CFAD_HUMAN	P07446 homo sapien
17	111	5.3	761	1	NETR_MOUSE	P08762 mus musculus
18	110	5.3	269	1	EL2B_HUMAN	P02128 homo sapien
19	110	5.3	492	1	FALO_BOVIN	P07443 bos taurus
20	108	5.2	593	1	CFA1_HUMAN	P0156 homo sapien
21	108	5.2	786	1	STUB_DRONE	Q03319 drosophila
22	107.5	5.2	686	1	MAS2_HUMAN	P00187 homo sapien
23	107	5.1	603	1	CPA1_MOUSE	Q61129 mus musculus
24	105.5	5.1	604	1	CFAL1_RAT	Q9nwv3 ratus norvegicus
25	105	5.0	269	1	EL2_BOVIN	P2461 bos taurus
26	105	5.0	274	1	TRV5_ANOGA	P3039 anophelis gambiae
27	104.5	5.0	268	1	TRV5_STRE	Q51179 streptomyces
28	103.5	5.0	271	1	EL2_RAT	P00774 ratus norvegicus
29	103.5	5.0	416	1	FA9_BOVIN	P0741 bos taurus
30	101.5	4.9	281	1	TRV5_DRORR	P5630 drosophila
31	99	4.8	855	1	ST14_MOUSE	P5677 mus musculus
32	98.5	4.7	260	1	COG5_HUMAN	P08897 hypodermes
33	98.5	4.7	267	1	TRV5_ANOGA	P3041 anophelis gambiae

DR PROSITE; PS00672; V8_HIS; 1.
 DR PROSITE; PS00673; V8_SER; 1.
 KW Hydrolase; Serine protease; signal.
 FT SIGNAL 1 ?
 FT PROPEP ? 94
 FT CHAIN 95 316 GLUTAMYL ENDopeptidase.
 FT ACT_SITE 141 141 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 261 261 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 126 142
 FT DISULFID 275 279
 SQ SEQUENCE 316 AA; 33611 MW; 96D7552CB7089B09 CRC64; 9.38; Score 194; DB 1; Length 316;
 Best Local Similarity 24.0%; Pred. No. 4.3e-09;
 Matches 79; Conservative 43; Mismatches 143; Indels 64; Gaps 15;

Query Match 9.38; Score 194; DB 1; Length 316;
 Best Local Similarity 24.0%; Pred. No. 4.3e-09;
 Matches 79; Conservative 43; Mismatches 143; Indels 64; Gaps 15;

CC MEDLINE=89108019; PubMed=3145906;
 RA Smith H., de Jong A., Bron S., Venema G.;
 RT "Characterization of signal-sequence-coding regions selected from the
 Bacillus subtilis chromosome.";
 RL Gene 70:351-361(1988).
 RN [4]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96118702; PubMed=7496533;
 RA Saxild H.H., Jacobsen J.H., Nygaard P.;
 RT "Functional analysis of the Bacillus subtilis purT gene encoding
 formate-dependent glycaminamide ribonucleotide transformylase.";
 RL Microbiology 141:2211-2218(1995).
 CC -!- SUBUNIT: MONOMER.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2B; ALSO KNOWN AS THE
 V8 FAMILY.
 CC -!- CAUTION: CALLED "METALLOPROTEASE", BUT CLEARLY BELONGS TO THE S2
 FAMILY OF SERINE PROTEASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 of the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to licensed@isb-sib.ch).
 CC

CC DR EMBL: 110505; AAA22604; 1; -.
 DR EMBL: AB006424; BAA3312; 1; -.
 DR EMBL: M22916; AAA22832; 1; -.
 DR EMBL: Z99105; CAB12018; 1; -.
 PIR: A35122; A35122.
 HSSP: P00756; 1SGF.
 MEROPS: S01_272; -.
 SubList; BG1690; mpr.
 InterPro; IPR000128; Ser_proteas_v8.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PRO0089; trypsin; 1.
 DR SMART: SM00020; tryp_SPC; 1.
 DR PROSITE: PS00672; V8_HIS; 1.
 DR PROSITE; PS00673; V8_SER; 1.
 DR KW Hydrolase; Serine protease; signal; zymogen; Complete proteome.
 FT SIGNAL 1 34 AA; 33842 MW; D41788E8D52AE94 CRC64; 6.58; Score 135.5; DB 1; Length 313;
 FT PROPEP 35 93 EXTRACELLULAR METALLOPROTEASE.
 FT CHAIN 94 313 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 146 146 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 267 267 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 131 131 BY SIMILARITY.
 FT CONFLICT 61 68 QVSAPYEG -> PLESTAOA (IN REF. 3).
 SQ SEQUENCE 313 AA; 33842 MW; D41788E8D52AE94 CRC64;

Query Match 6.58; Score 135.5; DB 1; Length 313;
 Best Local Similarity 22.2%; Pred. No. 0.00035;
 Matches 77; Conservative 38; Mismatches 121; Indels 111; Gaps 17;

CC 72 GTPPTPYEAKQYLSYELYANGSRTEQGYIILSSSD--GAQHRSQGSSGSRRQ 129
 FT ACT SITE 146 146 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 267 267 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 131 131 BY SIMILARITY.
 FT CONFLICT 61 68 QVSAPYEG -> PLESTAOA (IN REF. 3).
 DB 76 LYQQTELEKNIQLQPSIIGDERTTISSTSPYRATVQSLVYKNTSSYGCQFL 135

Qy 165 VAEKHLTAACIH-----DGKTYVKSTOKLVRGFLPKPKFDGRRGAND 208
 Db 136 VNPATVWVACHCVYQDHWASTITAAPGRNQSSYPTV-----175

Qy 209 STSAMPEQMKFQHNRVKRTHVPGWIKGN-ANDIGMDYDVALLEKKPHKKRPMKGV 265
 Db 176 -SGTMFYSVK-GWTESKDNYDGAIKLNGSPNTVGNW-YGVRTTNSSSP-----VGL 225

CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=GP41;
 RX MEDLINE=20130256; PubMed=2105291;
 RA Sloma A., Rudolph C.F., Rufo G.A. Jr., Sullivan B.J., Theriault K.A.,
 Ally D., Pero J.;
 RT "Gene encoding a novel extracellular metalloprotease in Bacillus
 subtilis";
 RT subtilis";
 RL J. Bacteriol. 172:1024-1029(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.;
 RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
 Bacillus subtilis chromosome.";
 RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RP SEQUENCE OF 1-68 FROM N.A.

Result No.	Score	Query Match	Length	DB ID	Description
1	194	9.3	316	2 A45134	endopeptidase (EC metalloproteinase
2	135	6.5	313	2 A35122	coagulation factor
3	120	5.8	482	1 EXRT	pancreatic elastase
4	119	5.7	269	2 A26823	pancreatic protease
5	117	5.6	522	2 T29167	pancreatic elastase
6	115	5.6	269	2 B26823	pancreatic elastase
7	115	5.6	271	2 A25528	pancreatic elastase
8	115	5.5	258	2 S70439	probable pancreatic
9	115	5.5	267	4 A56615	trypsin-like protease
10	114	5.5	238	1 TRN5Y	complement factor
11	114	5.5	246	1 DBHU	pancreatic elastase
12	113	5.4	266	1 ELETT1	pancreatic elastase
13	113	5.4	266	1 ELPG	trypsin (EC 3.4.21)
14	112	5.4	259	1 TRNSG	hypothetical protease
15	112	5.4	273	2 E85765	probable secreted
16	112	5.4	273	2 H64915	complement factor
17	111	5.4	488	1 EXHU	coagulation factor
18	111	5.3	761	2 JC3759	brain-specific serine
19	110	5.3	269	2 C26823	pancreatic elastase
20	110	5.3	492	1 EXBO	coagulation factor
21	110	5.3	1582	2 T13108	hypothetical protease
22	109	5.3	405	2 T35117	probable secreted
23	108	5.2	583	2 A29154	complement factor
24	108	5.2	786	1 A41547	serine proteinase
25	108	5.2	1047	2 A55617	masquerade precursors
26	107	5.2	236	2 A25566	T-cell suppressor
27	107	5.2	686	1 A59271	Reactive factor
28	105	5.0	274	2 S40004	trypsin-related protease
29	105	5.0	271	1 ELRT2	pancreatic elastase

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:17:56 ; Search time 48.03 Seconds (without alignments)

607.430 Million cell updates/sec

Title: US-09-072-384-18

Perfect score: 2080

Sequence: 1 MAGIPGLLFLLELLCAVQQ.....LKYAQICYWIKGNYLDREG 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	9.3	316	2 A45134	endopeptidase (EC metalloproteinase
2	135	6.5	313	2 A35122	coagulation factor
3	120	5.8	482	1 EXRT	pancreatic elastase
4	119	5.7	269	2 A26823	pancreatic protease
5	117	5.6	522	2 T29167	pancreatic elastase
6	115	5.6	269	2 B26823	pancreatic elastase
7	115	5.6	271	2 A25528	pancreatic elastase
8	115	5.5	258	2 S70439	probable pancreatic
9	115	5.5	267	4 A56615	trypsin-like protease
10	114	5.5	238	1 TRN5Y	complement factor
11	114	5.5	246	1 DBHU	pancreatic elastase
12	113	5.4	266	1 ELETT1	pancreatic elastase
13	113	5.4	266	1 ELPG	trypsin (EC 3.4.21)
14	112	5.4	259	1 TRNSG	hypothetical protease
15	112	5.4	273	2 E85765	probable secreted
16	112	5.4	273	2 H64915	complement factor
17	111	5.4	488	1 EXHU	coagulation factor
18	111	5.3	761	2 JC3759	brain-specific serine
19	110	5.3	269	2 C26823	pancreatic elastase
20	110	5.3	492	1 EXBO	coagulation factor
21	110	5.3	1582	2 T13108	hypothetical protease
22	109	5.3	405	2 T35117	probable secreted
23	108	5.2	583	2 A29154	complement factor
24	108	5.2	786	1 A41547	serine proteinase
25	108	5.2	1047	2 A55617	masquerade precursors
26	107	5.2	236	2 A25566	T-cell suppressor
27	107	5.2	686	1 A59271	Reactive factor
28	105	5.0	274	2 S40004	trypsin-related protease
29	105	5.0	271	1 ELRT2	pancreatic elastase

ALIGNMENTS

RESULT 1									
A45134									
endopeptidase (EC 3.4.-.-.), glutamate-specific - <i>Bacillus licheniformis</i>									
C:Species: <i>Bacillus licheniformis</i>									
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999									
C:Accession: A45134; S23078									
R:Kakudo, S.; Kiuchi, N.; Kitadokoro, K.; Fujiwara, T.; Nakamura, E.; Okamoto, H.; S: J. Biol. Chem. 267, 23782-23788 (1992)									
A:Title: Purification, characterization, cloning, and expression of a glutamic acid-specific endopeptidase									
A:Reference number: S23078; MUID:9215199									
A:Status: preliminary									
A: Molecule type: DNA									
A: Residues: 1-316 <KAK>									
A: Cross-references: GB:D10060; NID:9216263; PID:BA00949.1; PID:d1001415; PID:g21626									
A:Experimental source: ARCC 14580									
A:Note: sequence extracted from NCBI backbone (NCBIN:118784, NCBIPI:118785)									
R:Svendsen, I.; Breddam, K.									
Eur. J. Biochem. 204, 165-171, 1992									
A:Title: Isolation and amino acid sequence of a glutamic acid specific endopeptidase									
A:Reference number: S23078; MUID:9215199									
A:Status: preliminary									
A: Molecule type: protein									
A: Residues: 1-316 <SVE>									
C:Keywords: hydrolase									
Query Match									
Best Local Similarity 24.0%									
Pred. No. 1-316 <KAK>									
A:Cross-references: GB:D10060; NID:9216263; PID:BA00949.1; PID:d1001415; PID:g21626									
A:Experimental source: ARCC 14580									
A:Note: sequence extracted from NCBI backbone (NCBIN:118784, NCBIPI:118785)									
R:Svendsen, I.; Breddam, K.									
Eur. J. Biochem. 204, 165-171, 1992									
A:Title: Isolation and amino acid sequence of a glutamic acid specific endopeptidase									
A:Reference number: S23078; MUID:9215199									
A:Status: preliminary									
A: Molecule type: DNA									
A: Residues: 1-316 <KAK>									
A: Cross-references: GB:D10060; NID:9216263; PID:BA00949.1; PID:d1001415; PID:g21626									
A:Experimental source: ARCC 14580									
A:Note: sequence extracted from NCBI backbone (NCBIN:118784, NCBIPI:118785)									
R:Svendsen, I.; Breddam, K.									
Eur. J. Biochem. 204, 165-171, 1992									
A:Title: Isolation and amino acid sequence of a glutamic acid specific endopeptidase									
A:Reference number: S23078; MUID:9215199									
A:Status: preliminary									
A: Molecule type: protein									
A: Residues: 1-316 <SVE>									
C:Keywords: hydrolase									
Query Match									
Best Local Similarity 24.0%									
Pred. No. 1-316 <KAK>									
A:Cross-references: GB:D10060; NID:9216263; PID:BA00949.1; PID:d1001415; PID:g21626									
A:Experimental source: ARCC 14580									
A:Note: sequence extracted from NCBI backbone (NCBIN:118784, NCBIPI:118785)									
R:Svendsen, I.; Breddam, K.									
Eur. J. Biochem. 204, 165-171, 1992									
A:Title: Isolation and amino acid sequence of a glutamic acid specific endopeptidase									
A:Reference number: S23078; MUID:9215199									
A:Status: preliminary									
A: Molecule type: DNA									
A: Residues: 1-316 <KAK>									

QY 344 QWDMNGSPQDENVAVIRTPPLKYAQICYW 372
 Db 292 -----SSYNQTRTRKVEFNLNTNN 311

RESULT 2

A35122 metallocroteinase (EC 3.4.22.2) mpr precursor, extracellular - *Bacillus subtilis*
 C:Species: *Bacillus subtilis*
 C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 15-Oct-1999
 R:Accession: A35122; 140010; A69660
 R:Cross-references: GB:L10505; PIDN:AAA22604.1; PID:g143210; GB:M29036
 R:Smith, H.; de Jong, A.; Bron, S.; Verema, G.
 A:Title: Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*.
 A:Reference number: A35122; MUID:00130256
 A:Accession: A35122
 A:Status: preliminary
 A: Molecule type: DNA
 A:Accession: 1-313 <SL0>
 A:Cross-references: GB:L10505; PIDN:AAA22604.1; PID:g143210; GB:M29036
 A:Title: Characterization of signal-sequence-coding regions selected from the *Bacillus* s
 A:Reference number: 139994; MUID:89108019
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A:Accession: 1-60,65,'L',67,'S',69,'AQA'<RE3>
 A:Cross-references: GB:M22916; PID:9143702
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Bertein
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997

A:Authors: Fouger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 leich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hull, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A.; Authors: Lauber, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel
 Y., M.; Ogawa, K.; Ogihara, A.; Oniega, B.; Park, S.H.; Patro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A.; Authors: Schleich, S.; Schrothe, R.; Scrofone, P.; Sekiguchi, J.; Sekowska, A.; Seror
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshihara, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69380; MUID:98044033
 A:Accession: A69660
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A:Accession: 1-313 <KUN>
 A:Cross-references: GB:AL009126; PID:92632457; PID:el1182176;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: mpr
 C:Keywords: hydrolase

Query Match 6.5% Score 135.5; DB 2; Length 313;
 Best Local Similarity 22.28; Pred. No. 0.00091;
 Matches 77; Conservative 38; Mismatches 121; Indels 111; Gaps 17;

QY 72 GTPPIPYEEAKQYQISYELYANGSRTETQVGYIYLSSGGD-GAQHRDSSGSKSRKKQ 129
 Db 29 GVPAKAAENPQTSVSNTGKEADATKNOT-----SKADQVSAPYETGKTEKS----- 75

QY 130 IGYDSRF-----SIFGKD-----FLINYPFESTVSLST-----GCCTGTL 164
 Db 76 LYGGTQELEKNQTQLPSSIIQGDERTRISSTTSFPPYRQLQSIKYKPNTSSTYGCIGFL 135

QY 165 VAERKHVLTAAHCH-----DGKTYVKGTKQKLRVGFKLKPKFKDGGGRAND 208

QY 136 VNPPTVVTAGCHYQISQDHGWASTITAAGPRNGSYPVY----- 175

Db 209 STSAMPEQMKEFWIRVKRTHPKGWIKGN--ANDIGMDYDVALLEKKPKHFKMKGIV 265

Db 176 -SGTMFYSVK-GWTESKDNYGAIKLNQSPGNTVGW-YGYRTNSSLSS-----vGL 225
 QY 266 SPPAKOLPQGRHISGYDNDRPGNLVYRCQDADQPGASGSYVVRMW 325
 Db 226 SSSVTGFPCCDKTFCGTMWSDTKPIR-----SAETYKLT-YTDYGCOSGPVY---- 272
 QY 326 KROQOKWERRLIGTFSGHQWDMNGSPQDENAVAYTRPLKYAQICYW 372
 Db 273 -RNYSDTGTQTAIHT-----NGG-SSSYNLGTRVTFNNIQYW 310

RESULT 3

EXRT coagulation factor Xa (EC 3.4.21.6) precursor - rat
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Accession: 07-Jan-1995 #sequence_change 07-Feb-1997 #text_change 08-Dec-2000
 C:Accession: S49075; JCA670; PS0190; I62745
 R:Stanton, C.; Ross, P.; Huston, S.; Wallin, R.
 Thromb. Res. 80, 63-73, 1995
 A:Title: Evidence for competition between vitamin K-dependent clotting factors
 A:Reference number: A58498; MUID:36093366
 A:Accession: S49075
 A:Molecule type: mRNA
 A:Residues: 1-482 <STAL>
 A:Cross-references: EMBL:X79807; PIDN:CAA56202.1; PID:9506601
 A:Note: submitted to the EMBL Data Library, June 1994
 A:Note: neither the complete nucleic acid sequence nor the complete translation are s
 R:Stanton, C.; Ross, P.; Huston, S.; Wallin, R.
 R:Enjoji, K.; Mizazaki, K.; Kato, H.
 J. Biochem. 109, 890-898, 1991
 A:Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat
 A:Reference number: PS0190; MUID:32041742
 A:Accession: PS0191
 A:Molecule type: protein
 A:Residues: 41-58 'X', 60-65 <ENJ1>
 A:Accession: JC4670; MUID:96194815
 A:Accession: JC4670
 A:Molecule type: mRNA
 A:Residues: 1-482 <STAL>
 A:Cross-references: EMBL:X79807; PIDN:CAA56202.1; PID:9506601
 A:Experimental source: Cos-1 cell
 A:Accession: S49075; JCA670; PS0190; I62745
 A:Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat
 A:Reference number: PS0190; MUID:32041742
 A:Accession: PS0191
 A:Molecule type: protein
 A:Residues: 41-58 'X', 60-65 <ENJ1>
 A:Accession: PS0190
 A:Molecule type: protein
 A:Residues: 183-186 'X', 188-207 <ENJ2>
 R:Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Niho, Y.
 Eur. J. Haematol. 52, 162-168, 1994
 A:Title: Primary structure of the partial nucleotide sequences and deduced primary structure
 A:Reference number: I46196; MUID:94222160
 A:Accession: I46196
 A:Accession: I62745
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Accession: GB:D21215; MUID:9415309; PIDN:BAA04756.1; PID:9455396
 A:Cross-references: GB:D21215
 C:Function:
 A:Description: catalyzes the proteolytic activation of prothrombin to thrombin in the
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homolog
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-40/Domain: propeptide #status predicted <PRO>
 F:25-84/Domain: Gla domain homology <Gla>
 F:41-179/Domain: coagulation factor X light chain #status predicted <LCH>
 F:90-121/Domain: EGF homology <EG1>
 F:129-164/Domain: EGF homology <EG2>
 F:183-482/Domain: coagulation factor X heavy chain #status predicted <HCH>
 F:183-231/Domain: activation peptide #status predicted <APN>
 F:231-482/Domain: coagulation factor X heavy chain #status predicted <ACT>
 F:22-460/Domain: trypsin homology <TRY>
 F:45-47 54, 56, 59, 60, 65, 66, 69, 72, 79/Modified site: gamma-carboxyglutamic acid (Glu) #
 F:57-62, 90-101, 95-110, 112-121, 129-140, 136-149, 151-164, 172-340, 238-243, 259-275, 388-402

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:25:48 ; Search time 84.84 Seconds

(without alignments)
675.846 Million cell updates/sec

Title: US-09-072-384-15

Perfect score: 2131

Sequence: 1 MAGIPGLLFLLEFLCAVQGQ.....TKGNYLDCREGDTVFPFGSN 392

Scoring table: BLOSUM62

Gapext 0.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_17:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2080	97.6	383	4	095084	095084 PRELIMINARY; ID: 095084; AC: 095084; DT: 01-MAY-1999 (TREMBLrel. 10, Created) 01-MAY-1999 (TREMBLrel. 10, Last sequence update) 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
2	1884.5	88.4	382	11	Q9D6X6	DE SERINE PROTEASE (HYPOTHETICAL 4.0 KDA PROTEIN) (PROTEASE, SERINE, 23).
3	1042	48.9	413	4	Q9BQP6	GN 2SIG13 OR DKFZP586B0719.
4	129.5	6.3	799	11	Q9DBI0	OS Homo sapiens (Human).
5	129.5	6.1	303	2	Q9EXR9	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
6	127.5	6.1	469	6	Q9GMD9	RT "A novel serine protease from human umbilical vein endothelial cells." NCBI-TaxID-9606;
7	127.5	6.0	1322	5	Q9NATO	RN [1] Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
8	126.5	5.9	678	11	Q9JJS8	RP SEQUENCE FROM N.A.
9	124.5	5.8	1322	5	Q9NJS5	RP TISSUE-UMBILICAL VEIN;
10	124	5.8	266	6	Q96444	RA Sheppard P., Blumberg H., Jelinek L., Foster D., O'Hara P.; Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
11	123.5	5.8	339	11	Q9QX91	RA TISSUE-UTERUS;
12	123.5	5.8	366	11	Q9QX85	RA Wambutt R., Heubner D., Newes H.W., Gassenhuber J., Wiemann S.; Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
13	123.5	5.8	541	11	Q9QX90	RA TISSUE-CERVIX CARCINOMA;
14	123.5	5.8	623	11	Q9JJP3	RA Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
15	123.5	5.8	643	11	Q9QX84	RA CC [-] SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
16	122	5.7	259	5	Q9XY61	DR EMBL; AF015287; AAD01553.1; [-]
17	122	5.7	449	5	Q9VYD8	DR EMBL; AL136914; CAB65848.1; [-]
18	120.5	5.7	482	11	Q63207	DR EMBL; BC001278; AAH01278.1; [-]
19	118	5.5	266	11	Q9D936	DR MEROPS; S01.309; [-]

ALIGNMENTS

20	118	5.5	573	5	Q9V516	09v516 drosophila
21	118	5.5	1376	5	Q9VQ8	09vq8 drosophila
22	117	5.5	258	5	Q9W5U8	09w5u8 drosophila
23	117	5.5	522	5	Q01771	001771 caenorhabdi
24	115.5	5.4	269	4	Q14243	014243 homo sapien
25	115	5.4	258	4	Q9UN11	Q9un11 homo sapien
26	114.5	5.4	279	5	Q96991	Q96991 manduca sex
27	113.5	5.3	418	5	Q9va87	Q9va87 drosophila
28	113.5	5.3	685	11	Q9z338	Q9z338 mus musculu
29	113.5	5.3	737	13	Q90422	090422 brachydanio
30	112.5	5.3	603	5	Q9VDV1	Q9vdv1 drosophila
31	112.5	5.3	686	4	Q9BZHO	Q9bzho homo sapien
32	112	5.3	249	13	Q9W7Q1	Q9w7q1 paralichthys
33	111.5	5.2	481	11	Q94740	Q94740 mus musculu
34	110	5.2	144.9	5	Q10922	Q10922 caenorhabdi
35	109.5	5.1	405	2	Q69973	Q69973 streptomyces
36	108.5	5.1	481	11	Q88947	Q88947 mus musculu
37	108.5	5.1	481	11	Q99132	Q99132 mus musculu
38	108	5.1	274	5	Q17086	Q17086 anopheles
39	108	5.1	750	13	Q9W633	Q9w633 cyprinus carpio
40	108	5.1	1047	5	Q24019	Q24019 drosophila
41	107.5	5.0	236	11	Q9Z1H1	Q9z1h1 mus musculu
42	107.5	5.0	427	10	Q9AYR4	Q9ayr4 chaetoceros
43	107.5	5.0	611	5	Q9VZS8	Q9vzs8 drosophila
44	106	5.0	460	5	Q9VFZ6	Q9vfz6 drosophila
45	106	5.0	762	13	Q9YIC6	Q9yic6 cyprinus carpio

DR	InterPro; IPRO01314; Chymotrypsin.	RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Saeki H., Sato K., Schoenbach C., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.-H., Weltz C., Whittaker C., Wilmung L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection.";
DR	Pfam; PF00089; trypsin; 1.	RA	
DR	PRTS; PR00722; CHYMOTRYPSIN.	RA	
DR	SMART; SM00202; TRYPSIN_SPc; 1.	RA	
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.	RA	
KW	Hydrolase; Protease; Serine protease.	RT	"Functional annotation of a full-length mouse cDNA collection.";
SEQUENCE	383 AA; 43001 MW; 466C11ABFD5E8F CRC64;	RL	Nature 409:695-690(2001).
CC	-1. SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).	CC	-1. SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR	EMBL; AK009847; BAB26501.1; -	DR	EMBL; AK009847; BAB26501.1; -
DR	MGI; MGI:1923703; 231046G15Rik.	DR	MGI; MGI:1923703; 231046G15Rik.
DR	InterPro; IPRO01314; Chymotrypsin.	DR	InterPro; IPRO01314; Chymotrypsin.
DR	InterPro; IPRO01254; Trypsin.	DR	InterPro; IPRO01254; Trypsin.
PFAM	PF00088; Trypsin; 1.	DR	PFAM; PF00088; Trypsin; 1.
PRNTS	PR00722; CHYMOTRYPSIN.	DR	PRNTS; PR00722; CHYMOTRYPSIN.
SMART	SM00002; TRYSPC; 1.	DR	SMART; SM00002; TRYSPC; 1.
PROSITE	PS00134; TRYPSIN_HIS; UNKNOWN_1.	DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW	Hydrolase; Serine protease.	DR	Hydrolase; Serine protease.
SEQUENCE	382 AA; 43053 MW; 69C9A7080E5B2306 CRC64;	SQ	SEQUENCE 382 AA; 43053 MW; 69C9A7080E5B2306 CRC64;
Query	Match 97.6%; Score 2080; DB 4; Length 383;	Query	Match 88.4%; Score 1884.5; DB 11; Length 382;
Best	Local Similarity 100.0%; Pred. No. 1. 8e-182;	Best	Local Similarity 90.6%; Pred. No. 1.e-164;
Local	Matches 383; Conservative 0; Mismatches 0; Gaps 0;	Local	Matches 20; Indels 1; Gaps 1;
Matches	347; Conservative 15; Mismatches 20; Gaps 1;	Matches	347; Conservative 15; Mismatches 20; Gaps 1;
Db	1 MAGIPGLLFLFLFLCAGQVSQPSYAPKPKTPAYRLPVVLPOSTLNLAKEPLPDEGAEAKLE 60	QY	1 MAGIPGLLFLFLFLCAGQVSQPSYAPKPKTPAYRLPVVLPOSTLNLAKEPLPDEGAEAKLE 60
Db	1 MAGIPGLLFLFLFLCAGQVSQPSYAPKPKTPAYRLPVVLPOSTLNLAKEPLPDEGAEAKLE 60	Db	1 MAGIPGLLFLFLFLCAGQVSQPSYAPKPKTPAYRLPVVLPOSTLNLAKEPLPDEGAEAKLE 60
QY	61 VSSCGPQCHKGTPPLPTEEAQKOLSYSTLYANGSRSTETQYIYLSSGDAQHRSGS 120	QY	61 VSSCGPQCHKGTPPLPTEEAQKOLSYSTLYANGSRSTETQYIYLSSGDAQHRSGS 120
Db	61 VSSCGPQCHKGTPPLPTEEAQKOLSYSTLYANGSRSTETQYIYLSSGDAQHRSGS 120	Db	61 VSSCGPQCHKGTPPLPTEEAQKOLSYSTLYANGSRSTETQYIYLSSGDAQHRSGS 120
QY	121 SGSRRKQIYGYDSRFSFSGKDFLNNPFESTSVKLSTRGCTGTVAAKHLTAACHTHDG 180	QY	121 SGSRRKQIYGYDSRFSFSGKDFLNNPFESTSVKLSTRGCTGTVAAKHLTAACHTHDG 180
Db	121 SGSRRKQIYGYDSRFSFSGKDFLNNPFESTSVKLSTRGCTGTVAAKHLTAACHTHDG 180	Db	120 TGRSRRKQIYGDGFEDLNNPFESTSVKLSTRGCTGTVAAKHLTAACHTHDG 179
QY	181 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 240	QY	181 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 240
Db	181 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 240	QY	181 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 240
QY	241 IGMDDYALLEKKPHKRFKEMIGVSPAKOLPGGRHIFSGYDNDRPGNLVYRFDYKDE 300	Db	180 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 239
Db	241 IGMDDYALLEKKPHKRFKEMIGVSPAKOLPGGRHIFSGYDNDRPGNLVYRFDYKDE 300	QY	180 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 239
QY	301 TYDLYQQCDQDAGQGASGSVYVNMWKROQQKWERKIGIFSGHONVDMNGSPQDENYAVR 360	Db	180 KTVKGKTOKLKVGLPKPKEFKDGRGANDSTSAMPQMKFQWIRVKRTHVPGKWNAND 239
Db	301 TYDLYQQCDQDAGQGASGSVYVNMWKROQQKWERKIGIFSGHONVDMNGSPQDENYAVR 360	QY	241 IGMDDYALLEKKPHKRFKEMIGVSPAKOLPGGRHIFSGYDNDRPGNLVYRFDYKDE 300
QY	361 ITPLKYAQICYWIKGNYLDCREG 383	Db	240 IGMDDYALLEKKPHKRFQFIKGNYLDCREG 383
Db	361 ITPLKYAQICYWIKGNYLDCREG 383	QY	360 ITPLKYAQICYWIKGNYLDCREG 383
RESULT	2	RESULT	3
Q9DQX6	PRELIMINARY;	Q9BQ6	PRELIMINARY;
ID	PRTR;	ID	PRTR;
AC	382 AA.	AC	413 AA.
Q9D6X6	PRELIMINARY;	Q9BQ6	PRELIMINARY;
DT	01-JUN-2001 (TREMBLrel. 17, Created)	DT	01-JUN-2001 (TREMBLrel. 17, Created)
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	231046G15RIK PROTEIN.	DE	231046G15RIK PROTEIN.
GN	231046G15RIK.	GN	231046G15RIK.
OS	Mus musculus (Mouse).	OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX	NCBI_TaxID=10090;	OX	NCBI_TaxID=9006;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=TONQUE;	RC	STRAIN=C57BL/6J; TISSUE=TONQUE;
RD	MEDLINE=21085666; PubMed=11217851;	RD	MEDLINE=21085666; PubMed=11217851;
RA	01-JUN-2001 (TREMBLrel. 17, Created)	RA	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
RA	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	RA	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
RA	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	RA	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
RA	231046G15RIK PROTEIN.	RA	231046G15RIK PROTEIN.
RA	231046G15RIK.	RA	231046G15RIK.
RA	Mus musculus (Mouse).	RA	Mus musculus (Mouse).
RA	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.	RA	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RA	NCBI_TaxID=10090;	RA	NCBI_TaxID=9006;
RA	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	SEQUENCE FROM N.A.	RC	SEQUENCE FROM N.A.
RD	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.	RD	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RA	AL121939; CAC35071.1;	RA	AL121939; CAC35071.1;
RA	818D9C951BD2D6C1 CRC64;	RA	818D9C951BD2D6C1 CRC64;

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

run on: December 2, 2001, 16:16:53 ; Search time 64.47 Seconds
 (without alignments)
 450.391 Million cell updates/second

title: US-09-072-384-15
 2131 1 MAGPGLFLFLFLLCAYGQ.....IKGNYLDREGDTVFPQGSN 392

target score:
 sequence:

scoring table: BLOSUM62

searched: 522463 seqs, 74073290 residues

total number of hits satisfying chosen parameters: 522463

minimum DB seq length: 0
 maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

database : A_Geneseq_1101.*

1: /SIDS2/gcdata/geneseq/geneseq/AA1980.DAT: *
 2: /SIDS2/gcdata/geneseq/geneseq/AA1981.DAT: *
 3: /SIDS2/gcdata/geneseq/geneseq/AA1982.DAT: *
 4: /SIDS2/gcdata/geneseq/geneseq/AA1983.DAT: *
 5: /SIDS2/gcdata/geneseq/geneseq/AA1984.DAT: *
 6: /SIDS2/gcdata/geneseq/geneseq/AA1985.DAT: *
 7: /SIDS2/gcdata/geneseq/geneseq/AA1986.DAT: *
 8: /SIDS2/gcdata/geneseq/geneseq/AA1987.DAT: *
 9: /SIDS2/gcdata/geneseq/geneseq/AA1988.DAT: *
 10: /SIDS2/gcdata/geneseq/geneseq/AA1989.DAT: *
 11: /SIDS2/gcdata/geneseq/geneseq/AA1990.DAT: *
 12: /SIDS2/gcdata/geneseq/geneseq/AA1991.DAT: *
 13: /SIDS2/gcdata/geneseq/geneseq/AA1992.DAT: *
 14: /SIDS2/gcdata/geneseq/geneseq/AA1993.DAT: *
 15: /SIDS2/gcdata/geneseq/geneseq/AA1994.DAT: *
 16: /SIDS2/gcdata/geneseq/geneseq/AA1995.DAT: *
 17: /SIDS2/gcdata/geneseq/geneseq/AA1996.DAT: *
 18: /SIDS2/gcdata/geneseq/geneseq/AA1997.DAT: *
 19: /SIDS2/gcdata/geneseq/geneseq/AA1998.DAT: *
 20: /SIDS2/gcdata/geneseq/geneseq/AA1999.DAT: *
 21: /SIDS2/gcdata/geneseq/geneseq/AA2000.DAT: *
 22: /SIDS2/gcdata/geneseq/geneseq/AA2001.DAT: *

summaries

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

result No.	Score	Query Length	DB ID	Description
1	2131	100.0	392 22 AA8973	Human Zsig13 vari
2	2085	97.8	392 22 AA8972	Human Zsig13 vari
3	2080	97.6	383 20 AA108657	Human transmembr
4	2080	97.6	383 20 AA098660	W09927094 Seq ID
5	2080	97.6	383 20 AA133390	Amino acid sequen
6	2080	97.6	383 21 AAB25592	protein encoded b
7	2080	97.6	383 21 AAB25618	protein encoded b
8	2080	97.6	383 21 AA88277	Human TANGO 186 p
9	2080	97.6	383 21 AA87270	Human signal pep
10	2080	97.6	383 21 AA53627	A bone marrow sec
11	2080	97.6	383 22 AA900257	Human op202 sec

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description
Zs1913 variant
Zs1913 variant
Human transmembrane protein encoded by
W0997094 Seq ID 1
Amino acid sequence of
Human TANGO-186 protein
Human signal peptide
A bone marrow secreted protein
Human protein

US6133420-A.

US6133420-A.

98US-0072384.

(ZYNO) ZYMOGENETICS I
Sheppard PO;
WPI: 2001-060090/07.
N-PSDB; AAC91783.

ALIGNMENTS

RESULT	1
AAB48973	
ID	AAB48973 standard; Protein: 392 AA.
XX	
AC	AAB48973;
XX	
DT	27-MAR-2001 (first entry)
XX	
DE	Human zsig13 variant #2, SEQ ID NO:15.
XX	
KW	Human zsig13; serine protease; chromosome 11q22.1; elastase homologue;
KW	glutamyl endopeptidase homologue; factor X homologue; trypsin homologue;
KW	trypsinogen homologue; mast cell protease homologue; collagenase homologue; protein degradation; food processing; brewing;
KW	alcohol production; laundry detergent component.

SUMMARIES							
Result No.	Score	Query Match	Length	DB ID	Description		
1	2131	100.0	392	22	AA848973	Human Zsig13 varia	PN XX
2	2085	97.8	392	22	AA848972	Human Zsig13 varia	PN XX
3	2080	97.6	383	20	AAY08657	Human transmembran	PN XX
4	2080	97.6	383	20	AAY08660	W09927094 Seq ID 1	PN XX
5	2080	97.6	383	20	AAY13390	Amino acid sequenc	PA XX
6	2080	97.6	383	21	AA825592	Protein encoded by	PA XX
7	2080	97.6	383	21	AA825618	Protein encoded by	PI PT
8	2080	97.6	383	21	AAY89277	Human TANGO 1B	DR DR
9	2080	97.6	383	21	AAV87270	Human signal pepti	DR DR
10	2080	97.6	383	21	AAV53627	A bone marrow secr	DR DR
11	2080	97.6	383	22	AA890258	Human DPC422 secr	DR DR

PT New isolated serine protease (designated zsig13), useful in industrial
 PT processes to degrade unwanted proteins or alter the characteristics of
 PT protein-containing composition, as well as in industrial applications
 PT (e.g. brewing) -

XX Claim 1; Column 35-38; 26pp; English.

XX The invention relates to human zsig13 proteins (AAB48972-B48874), and
 CC to DNA encoding them (AAC91782-C91784). The invention also relates to
 CC expression vectors and host cells comprising a human zsig13 DNA, and the
 CC recombinant production of a human zsig13 protein or its precursor.
 CC zsig13 is a serine protease, and has significant homology to Bacillus
 CC licheniformis glutamyl endopeptidase, human clotting factor X, human
 CC elastase, rat mast cell protease, Streptomyces trypsin, bovine
 CC trypsinogen, and Hypoderma lineatum collagenase. The gene encoding human
 CC zsig13 is located on chromosome 11q22.1. zsig13 is useful in industrial
 CC processes to degrade unwanted proteins or alter the characteristics of
 CC protein-containing compositions. It may also be used in industrial
 CC applications in which proteases are utilised, including food processing,
 CC brewing and alcohol production, and as a component of a laundry
 CC detergent. The present sequence represents a human zsig13 variant.

XX Sequence 392 AA.

Query	Match	Score	Length
Query	100 0%	Score 2131;	DB 22;
Best Local Matches	100 0%	Pred. No. 3.8e-152;	Length 392;
Local Similarity	100 0%	Mismatches 0;	Indels 0;
Matches	392;	Conservative	Gaps
Db	1 magipglflflffllcavgovspysapwptwpxylpvvlpqstlnlakpdfgneaakle 60		
Qy	1 MAGIPGLFLFLFFLLCAVGOSVSPYSAWPWTWPAVLPPVVLPPQSTLNLAKEPDFGEAEAKLE 60		
Db	1 vssscgpochkgtpalptyeakqlyseytlyangsrttgcgqytlsscdgqahrdsgs 120		
Qy	61 vssscgpochkgtpalptyeakqlyseytlyangsrttgcgqytlsscdgqahrdsgs 120		
Db	61 vssscgpochkgtpalptyeakqlyseytlyangsrttgcgqytlsscdgqahrdsgs 120		
Qy	121 SGKSRKRQTYGDSRFSRIFGKDFLNLNPFTSYKLSTGCTGTLYAEKHYLTAACIHGD 180		
Db	121 sgksrkqrqiyasrsfsigkdfllnlypistsvkstgtqivtaaikvhq 180		
Qy	181 KTYKKGTOQLRVGFLPKPKFDGGRCANDSTSAMPOMKWQWIRKRTHYPKGWIGKNGAND 240		
Db	181 ktykkgtoqlrvgflpkpkfdggcgandstsampeqmkgqiwirkrthypkgwlgkngand 240		
Qy	241 IGMDYDALLELKEPKHRKFEMKIGVSPAKQLPGRIRHESGYDNDRPGNLYVRFCDVKDE 300		
Db	241 igmdydallelkpkhrkmtkgvspakqlppgrihfegyndrprgnlyrfdvkd 300		
Qy	301 TYDLYQOCDAQGASGSVGVRMWKQROOKWERKINGIFSGHOMVDMINGSPQDFNNAVR 360		
Db	301 tydlyqocdaqprasgsgtvrmwkqrgqwerkimgifsghqwdmngspqdfnnavr 360		
Qy	361 ITPLKYAQIQCWIGNYLCREGDTVPPCSN 392		
Db	361 itpkyaqicwyikqnydcregdtvppcsn 392		
RESULT	2		
	AAB48972		
	ID AAB48972	standard; Protein: 392 AA.	
	XX		
	AC AAB48972;		
	XX		
	DT 27-MAR-2001	(first entry)	
	XX		
	DE Human Zsig13 variant #1, SEQ ID NO:2.		
	XX		
	KW Human Zsig13; serine protease; chromosome 11q22.1; elastase homologue; gilumitrope endopeptidase homologue; factor X homologue; trypsin homologue; trypsinogen homologue; mast cell protease homologue; collagenase homologue; protein degradation; food processing; brewing; collagenase, homologue;		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:15:34 ; Search time 35.78 Seconds
(without alignments)

240.882 million cell updates/sec

Title: US-09-072-384-18

Perfect score: 2080

Sequence: 1 MAGIPGLLFLLLFLCAVQQ.....LKYAQICYWIKGNYLDCREG 383

Scoring table: BLOSUM62

Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMBO.pep:*

2: /cgn2_6/ptodata/2/1aa/5B_COMBO.pep:*

3: /cgn2_6/ptodata/2/1aa/6A_COMBO.pep:*

4: /cgn2_6/ptodata/2/1aa/6B_COMBO.pep:*

5: /cgn2_6/ptodata/2/1aa/PCRTUS_COMBO.pep:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2080	100.0	383	4	US-09-072-384-18	Sequence 18, Appl
2	2080	100.0	392	4	US-09-072-384-15	Sequence 15, Appl
3	2044	98.3	392	4	US-09-072-384-1	Sequence 1, Appl
4	177	8.5	222	1	US-08-090-048-1	Sequence 1, Appl
5	177	8.5	222	2	US-08-292-550-1	Sequence 1, Appl
6	177	8.5	222	1	US-17-922-661A-1	Sequence 1, Appl
7	122	5.9	256	3	US-08-906-769-89	Sequence 89, Appl
8	122	5.9	256	3	US-08-906-616-89	Sequence 89, Appl
9	122	5.9	256	4	US-08-811-795-89	Sequence 89, Appl
10	122	5.9	256	4	US-08-639-079A-89	Sequence 89, Appl
11	122	5.9	256	4	US-09-012-431-89	Sequence 89, Appl
12	122	5.9	256	4	US-09-012-215-32	Sequence 89, Appl
13	122	5.9	256	4	US-09-012-692-89	Sequence 89, Appl
14	122	5.9	256	4	US-08-906-613-89	Sequence 89, Appl
15	122	5.9	256	5	PCT-US95-14412A-89	Sequence 2, Appl
16	117	5.6	437	1	US-08-487-037-2	Sequence 59, Appl
17	115	5.6	241	4	US-08-908-483-59	Sequence 44, Appl
18	114	5.5	228	4	US-08-914-483-44	Patent No. 5223425
19	114	5.5	253	6	5223425-8	Patent No. 5223425
20	113	5.4	238	6	5223425-5	Patent No. 5223425
21	113	5.4	250	6	5223425-4	Patent No. 5223425
22	112.5	5.4	223	1	US-08-288-091-13	Sequence 13, Appl
23	112.5	5.4	223	1	US-08-433-859-13	Sequence 13, Appl
24	112.5	5.4	223	2	US-08-471-167-13	Sequence 13, Appl
25	112.5	5.4	223	2	US-08-422-816-13	Sequence 13, Appl
26	112.5	5.4	223	2	US-08-296-149-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-072-384-18

; Sequence 18, Application US/09072384

; Patent No. 615420

GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.

; STREET: 1201 Eastlake Avenue, East

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/072,384

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, Gary E.

; REGISTRATION NUMBER: 31,648

; REFERENCE/DOCKET NUMBER: 97-16C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-442-6673

; TELEX:

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 383 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; NAME/KEY: Signal Sequence

; LOCATION: 1...19

; OTHER INFORMATION:

; US-09-072-384-18

Query Match 100.0%; Score 2080; DB 4; Length 383;
 Best Local Similarity 100.0%; Pred. No. 5e-223; Gaps 0;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIPGLLFFLFLICAVQSPYSPAPWKPTWPAYRLPVVLPOSTLNIAKPDFGAEAKLE 60
 Db 1 MAGIPGLLFFLFLICAVQSPYSPAPWKPTWPAYRLPVVLPOSTLNIAKPDFGAEAKLE 60
 Qy 61 VSSSGPQCHKGTPKTPTYEAQKLYSEPTLYANGSRTEHQVYIYLSSSGDGAQHRDGS 120
 Db 61 VSSSGPQCHKGTPKTPTYEAQKLYSEPTLYANGSRTEHQVYIYLSSSGDGAQHRDGS 120
 Qy 121 SGKSRKRQIYGYDSRFSIPLGKDFLLNPFTSYKLSPGCTGTVLAEKHVAHCLHDG 180
 Db 121 SGKSRKRQIYGYDSRFSIPLGKDFLLNPFTSYKLSPGCTGTVLAEKHVAHCLHDG 180
 Qy 181 KTYVGTQKLRVGFUJKPKFDGGGRGANDSTSAMPQMKFQWIRVKRTHVPKGWTGKGNAND 240
 Db 181 KTYVGTQKLRVGFUJKPKFDGGGRGANDSTSAMPQMKFQWIRVKRTHVPKGWTGKGNAND 240
 Qy 241 GMDDYALLELKPKHREKFMIGVSPPAQKLPGRHIFSGYNDRPGNLYRFDYKDE 300
 Db 241 GMDDYALLELKPKHREKFMIGVSPPAQKLPGRHIFSGYNDRPGNLYRFDYKDE 300
 Qy 301 TYDLYQQCDAQPGASGSQGYVYRMWKRQOQKWERKIGIFGSQHWDNNGSPQDENYAVR 360
 Db 301 TYDLYQQCDAQPGASGSQGYVYRMWKRQOQKWERKIGIFGSQHWDNNGSPQDENYAVR 360
 Qy 361 ITPLKYAQICYWIKGNYLDREG 383
 Db 361 ITPLKYAQICYWIKGNYLDREG 383

RESULT 2
 US-09-072-384-15
 ; Sequence 15, Application US/09072384
 ; Patent No. 6153420
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; TITLE OF INVENTION: AND MATERIALS AND METHODS FOR MAKING THEM
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/072,384
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PAKER, GARY E
 ; REGISTRATION NUMBER: 31,648
 ; REFERENCE/DOCKET NUMBER: 97-16C1
 ; TELEPHONE: 206-442-6673
 ; TELEX: 206-442-6678
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 392 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 NAME/KEY: Signal Sequence
 LOCATION: 1..19
 OTHER INFORMATION:
 US-09-072-384-15

Query Match 100.0%; Score 2080; DB 4; Length 392;
 Best Local Similarity 100.0%; Pred. No. 5.2e-223;
 Mismatches 0; Indels 0; Gaps 0;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIPGLLFFLFLICAVQSPYSPAPWKPTWPAYRLPVVLPOSTLNIAKPDFGAEAKLE 60
 Db 1 MAGIPGLLFFLFLICAVQSPYSPAPWKPTWPAYRLPVVLPOSTLNIAKPDFGAEAKLE 60
 Qy 61 VSSSGPQCHKGTPKTPTYEAQKLYSEPTLYANGSRTEHQVYIYLSSSGDGAQHRDGS 120
 Db 61 VSSSGPQCHKGTPKTPTYEAQKLYSEPTLYANGSRTEHQVYIYLSSSGDGAQHRDGS 120
 Qy 61 VSSSGPQCHKGTPKTPTYEAQKLYSEPTLYANGSRTEHQVYIYLSSSGDGAQHRDGS 120
 Db 61 VSSSGPQCHKGTPKTPTYEAQKLYSEPTLYANGSRTEHQVYIYLSSSGDGAQHRDGS 120
 Qy 121 SGKSRKRQIYGYDSRFSIPLGKDFLLNPFTSYKLSPGCTGTVLAEKHVAHCLHDG 180
 Db 121 SGKSRKRQIYGYDSRFSIPLGKDFLLNPFTSYKLSPGCTGTVLAEKHVAHCLHDG 180
 Qy 121 SGKSRKRQIYGYDSRFSIPLGKDFLLNPFTSYKLSPGCTGTVLAEKHVAHCLHDG 180
 Db 121 SGKSRKRQIYGYDSRFSIPLGKDFLLNPFTSYKLSPGCTGTVLAEKHVAHCLHDG 180
 Qy 181 KTYVGTQKLRVGFUJKPKFDGGGRGANDSTSAMPQMKFQWIRVKRTHVPKGWTGKGNAND 240
 Db 181 KTYVGTQKLRVGFUJKPKFDGGGRGANDSTSAMPQMKFQWIRVKRTHVPKGWTGKGNAND 240
 Qy 241 GMDDYALLELKPKHREKFMIGVSPPAQKLPGRHIFSGYNDRPGNLYRFDYKDE 300
 Db 241 GMDDYALLELKPKHREKFMIGVSPPAQKLPGRHIFSGYNDRPGNLYRFDYKDE 300
 Qy 301 TYDLYQQCDAQPGASGSQGYVYRMWKRQOQKWERKIGIFGSQHWDNNGSPQDENYAVR 360
 Db 301 TYDLYQQCDAQPGASGSQGYVYRMWKRQOQKWERKIGIFGSQHWDNNGSPQDENYAVR 360
 Qy 361 ITPLKYAQICYWIKGNYLDREG 383
 Db 361 ITPLKYAQICYWIKGNYLDREG 383

RESULT 3
 US-09-072-384-2
 ; Sequence 2, Application US/09072384
 ; Patent No. 6153420
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/072,384
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PAKER, GARY E
 ; REGISTRATION NUMBER: 31,648
 ; REFERENCE/DOCKET NUMBER: 97-16C1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-442-6673
 ; TELEX: 206-442-6678
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 392 amino acids

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:16:53 ; Search time 64.47 Seconds
(without alignments)
440.051 Million cell updates/sec

Title: US-09-072-384-18
Perfect score: 2050

Sequence: 1 MAGIPGULFLFLCAVGO LKYAQICYWIKGNYLDCREG 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101;*

1: /SIDS2/gcdata/geneseq/geneseq/AA1980.DAT;*
2: /SIDS2/gcdata/geneseq/geneseq/AA1981.DAT;*
3: /SIDS2/gcdata/geneseq/geneseq/AA1982.DAT;*
4: /SIDS2/gcdata/geneseq/geneseq/AA1983.DAT;*
5: /SIDS2/gcdata/geneseq/geneseq/AA1984.DAT;*
6: /SIDS2/gcdata/geneseq/geneseq/AA1985.DAT;*
7: /SIDS2/gcdata/geneseq/geneseq/AA1986.DAT;*
8: /SIDS2/gcdata/geneseq/geneseq/AA1987.DAT;*
9: /SIDS2/gcdata/geneseq/geneseq/AA1988.DAT;*
10: /SIDS2/gcdata/geneseq/geneseq/AA1989.DAT;*
11: /SIDS2/gcdata/geneseq/geneseq/AA1990.DAT;*
12: /SIDS2/gcdata/geneseq/geneseq/AA1991.DAT;*
13: /SIDS2/gcdata/geneseq/geneseq/AA1992.DAT;*
14: /SIDS2/gcdata/geneseq/geneseq/AA1993.DAT;*
15: /SIDS2/gcdata/geneseq/geneseq/AA1994.DAT;*
16: /SIDS2/gcdata/geneseq/geneseq/AA1995.DAT;*
17: /SIDS2/gcdata/geneseq/geneseq/AA1996.DAT;*
18: /SIDS2/gcdata/geneseq/geneseq/AA1997.DAT;*
19: /SIDS2/gcdata/geneseq/geneseq/AA1998.DAT;*
20: /SIDS2/gcdata/geneseq/geneseq/AA1999.DAT;*
21: /SIDS2/gcdata/geneseq/geneseq/AA2000.DAT;*
22: /SIDS2/gcdata/geneseq/geneseq/AA2001.DAT;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
AAY08657 1
ID AAY08657 standard; protein; 383 AA.
XX
AC AAY08657;
XX
XX 09-AUG-1999 (first entry)

XX Human transmembrane domain containing protein from clone HP10493.
DE
XX Transmembrane domain; human; nutrition; cytokine; cell differentiation; KW
immune stimulation; immune suppression; haemopoiesis; activin; KW
haemostatic; tissue growth; inhibin; chemostatic; chemokinetic; KW
thrombolytic; tumour inhibitor; anti-inflammatory; KW
gene therapy; screening.
XX Homo sapiens.
OS
XX W0992/094-A2.
PN
XX 03-JUN-1999.
PD
XX 20-NOV-1998; 98WO-JP05238.
PF
XX 25-NOV-1997; 97JP-0323129.
PR
XX (PROTEGENE INC.
PA (SAGA) SAGAMI CHEM RES CENT.
PA
XX Kato S, Kimura T, Sekine S;
XX WPI: 1999-357835/30.
DR N-PSDB; AAX77690, AAX77691.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2080	100.0	383	20 AAY08657	Human transmembran
2	2080	100.0	383	20 AAY08660	W0992/094 Seq ID 1
3	2080	100.0	383	20 AAY13390	Amino acid sequenc
4	2080	100.0	383	21 AAB25592	Protein encoded by
5	2080	100.0	383	21 AAB25618	Protein encoded by
6	2080	100.0	383	21 AAY88277	Human TANGO 186 Pr
7	2080	100.0	383	21 AAY87270	Human signal pepti
8	2080	100.0	383	21 AAY3627	A bone marrow secr
9	2080	100.0	383	22 AAB8058	Human PRO307 prote
10	2080	100.0	383	22 AAB48974	Human Zs1913 varia
11	2080	100.0	392	22 AAB48973	Human Zs1913 varia

PT Novel proteins containing transmembrane domains, useful as anti-inflammatory, immune stimulators/suppressors and tissue growth compounds
 PR XX
 PS Claim 1; Page 68-69; 89pp; English.
 XX

CC This invention describes novel human transmembrane containing proteins and their encoding nucleic acids. Although no specific use is given for the proteins, they may have a range of activities selected from nutritional uses, cytokine and cell differentiation, immune stimulation/suppression, haemopoiesis regulatory, tissue growth, activin/inhibin, chemostatic/chemokinetic, haemostatic/thrombolytic, receptor/ligand, tumour inhibitor, anti-inflammatory and other undefined activities. The cDNAs can be utilized as probes for gene diagnosis and expression of proteins. The transformed cells can be used for large scale detection of the corresponding ligands and for screening of novel low-molecular pharmaceuticals.
 XX

Sequence 383 AA;

Query Match 100.0%; Score 2080; DB 20; Length 383;
 Best Local Similarity 100.0%; Pred. No. 1e-146;
 Matches 383; Conservative 0; Mismatches 0; Gaps 0;
 Indels 0;

1 MAGIPGLLFLFLFLCAVGQVSPYSSAPMKPTPAYRPLPVVLPOSTLNLAKEPKPFGAEAKLE 60
 Db 1 magipglflflflcavgqvspsapmkptpayrplpvvlpostlnlaakepkpfgaeakle 60
 Qy 61 VSSSGPPOCHKGTPPLPTBEAKOYLSTETLYANGSRTEQYIYLSSGDSAQHRSQG 120
 Db 61 vssscgpochkgtpplptbeakoylstetlyangsrteqyiyllssgdaqhrsgs 120
 Qy 121 SGKSRKRQIYQYDSDRSFIEFGKDFLLNYPFSTSVKLSTGCTGTVAEKHVLTAACHTHDG 180
 Db 121 sgksrkrqiyqydsrsfiefgkdfllnypfstsvklistgtctgtvaekhvltaahchhdg 180
 Qy 181 KTVVKGTOKLRYGLKPKFKDGGRGANDSTSAMPQMKFQWVTKRTHPKWIKGNAND 240
 Db 181 ktvvkgtoklryglkpkfkdfggrrgandstsampqmkfqwvtrkrtphwkgnand 240
 Qy 241 IGMDDYDALLEKKPHRKEMKIGVSPPAKQPGGRTHFSGDNDRQGNLVYRFDVYKDE 300
 Db 241 igmddydalekkphrkemkigvspapkqpggrthfsgdndrqgnlvyrfdvkyde 300
 Qy 301 TYDLYQQCDAQPGASGSGVYRMWKROQKWERKIGIFSSHOWDMNGSSQDFNVAVR 360
 Db 301 tydlyqqcdaqpgasgsgvyrmwkroqkwerkigifsshowdmngssqdfnvavr 360
 Qy 361 ITPLKYAOTCYNKGNYLDCREG 383
 Db 361 itplkyadicywkgnyldcreg 383

RESULT 2
 ID AAY08660
 XX
 AAY08660 Standard; Protein; 383 AA.
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE WO9927094 Seq ID 12.
 XX
 Transmembrane domain; human; nutrition; cytokine; cell differentiation; immune stimulation; immune suppression; haemopoiesis; activin; regulatory tissue growth; inhibin; chemostatic; chemokinetic; haemostatic; thrombolytic; tumour inhibitor; anti-inflammatory; gene therapy; screening.
 XX
 Homo sapiens.
 OS

PN WO9927094-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 20-NOV-1998; 98WO-JP05238.
 XX
 PR 25-NOV-1997; 97JP-0323129.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 Kato S, Kimura T, Sekine S;
 PI
 DR WPI; 1999-357835/30.
 XX
 Novel proteins containing transmembrane domains, useful as anti-inflammatories, immune stimulators/suppressors and tissue growth compounds
 XX
 Disclosure; Page 87-89; 89pp; English.
 XX
 This invention describes novel human transmembrane containing proteins and their encoding nucleic acids. Although no specific use is given for the proteins, they may have a range of activities selected from nutritional uses, cytokine and cell differentiation, immune stimulation/suppression, haemopoiesis regulatory, tissue growth, activin/inhibin, chemostatic/chemokinetic, haemostatic/thrombolytic, receptor/ligand, tumour inhibitor, anti-inflammatory and other undefined activities. The cDNAs can be utilized as probes for gene diagnosis and as gene sources for gene therapy. The cDNAs can also be used for large scale expression of proteins. The transformed cells can be used for detection of the corresponding ligands and for screening of novel low-molecular pharmaceuticals.
 XX
 Sequence 383 AA;

Query Match 100.0%; Score 2080; DB 20; Length 383;
 Best Local Similarity 100.0%; Pred. No. 1e-146;
 Matches 383; Conservative 0; Mismatches 0; Gaps 0;
 Indels 0;

Qy 1 MAGIPGLLFLFLCAVGQVSPYSSAPMKPTPAYRPLPVVLPOSTLNLAKEPKPFGAEAKLE 60
 Db 1 magipglflflcavgqvspsapmkptpayrplpvvlpostlnlaakepkpfgaeakle 60
 Qy 61 VSSSGPPOCHKGTPPLPTBEAKOYLSTETLYANGSRTEQYIYLSSGDSAQHRSQG 120
 Db 61 vssscgpochkgtpplptbeakoylstetlyangsrteqyiyllssgdaqhrsgs 120
 Qy 121 SGKSRKRQIYQYDSDRSFIEFGKDFLLNYPFSTSVKLSTGCTGTVAEKHVLTAACHTHDG 180
 Db 121 sgksrkrqiyqydsrsfiefgkdfllnypfstsvklistgtctgtvaekhvltaahchhdg 180
 Qy 181 KTVVKGTOKLRYGLKPKFKDGGRGANDSTSAMPQMKFQWVTKRTHPKWIKGNAND 240
 Db 181 ktvvkgtoklryglkpkfkdfggrrgandstsampqmkfqwvtrkrtphwkgnand 240
 Qy 241 IGMDDYDALLEKKPHRKEMKIGVSPPAKQPGGRTHFSGDNDRQGNLVYRFDVYKDE 300
 Db 241 igmddydalekkphrkemkigvspapkqpggrthfsgdndrqgnlvyrfdvkyde 300
 Qy 301 TYDLYQQCDAQPGASGSGVYRMWKROQKWERKIGIFSSHOWDMNGSSQDFNVAVR 360
 Db 301 tydlyqqcdaqpgasgsgvyrmwkroqkwerkigifsshowdmngssqdfnvavr 360
 Qy 361 ITPLKYAOTCYNKGNYLDCREG 383
 Db 361 itplkyadicywkgnyldcreg 383

RESULT 2
 ID AAY13390
 XX
 AAY13390 standard; Protein; 383 AA.
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE WO9927094 Seq ID 12.
 XX
 Transmembrane domain; human; nutrition; cytokine; cell differentiation; immune stimulation; immune suppression; haemopoiesis; activin; regulatory tissue growth; inhibin; chemostatic; chemokinetic; haemostatic; thrombolytic; tumour inhibitor; anti-inflammatory; gene therapy; screening.
 XX
 Homo sapiens.
 OS

RESULT 3
 ID AAY13390

AAV13390 standard; Protein; 383 AA.

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

Q90OM protein - protein search, using sw model

Run on: December 2, 2001, 16:25:50 ; Search time 84.84 Seconds
 (without alignments)
 650,329 Million cell updates/second

Title: US-09-072-384-18

Perfect score: 2080.

Sequence: 1 MAGPGLLFLFLCAVGQ.LKYAQICYWIKGNYLDCREG 383

Scoring table: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0⁸
 Maximum Match 0⁸

```

SPTREMBL:*
 1: sp_archea:*
 2: sp_bacteria:*
 3: sp_fungi:*
 4: sp_human:*
 5: sp_invertebrate:*
 6: sp_mammal:*
 7: sp_mhc:*
 8: sp_organelle:*
 9: sp_phage:*
10: sp_plant:*
11: sp_rhodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

```

PreP. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID
1	2080	100.0	383	4 C95084
2	1884.5	90.6	382	11 Q9D6X6
3	1042	50.1	413	4 QBOPB6
4	1335.5	6.4	799	11 Q9DB10
5	1295.5	6.2	303	2 QBEXR9
6	129	6.2	469	6 Q9CMD9
7	1275	6.1	1322	5 Q9NATO
8	126.5	6.1	678	11 Q9JJ58
9	124.5	6.0	1322	5 Q9NJS5
10	124	6.0	266	6 Q96644
11	123.5	5.9	339	11 Q9QX91
12	123.5	5.9	366	11 Q9QX85
13	123	5.9	541	11 Q9QX90
14	123.5	5.9	623	11 Q9JJP3
15	123.5	5.9	643	11 Q9QX84
16	122	5.9	259	5 Q9YX61
17	122	5.9	449	5 Q9VDU8
18	120.5	5.8	482	11 Q632207
19	118	5.7	1376	5 Q9YOR8

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	383 AA.
095084	095084			
	1D			
	AC	095084;		
	DT	01-MAY-1999	(TREMBLrel. 10, Created)	
	DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
	DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
	DT		SERINE PROTEASE (HYPOTHETICAL 43.0 KDA PROTEIN)	
	DE		SERINE PROTEASE (HYPOTHETICAL 43.0 KDA PROTEIN) (PROTEASE, SERINE PROTEASE)	

GN ZSIG13, OR DKFZPP56B0719.
Eukaryota: Metazoa: Human: Chordata: Craniata: Vertebrates: Euteleostomi:
OC

NCBI_TaxID=9606;
OX
RN
[1]
RP
SEQUENCE FROM N.A.
RC
TISSUE=UMBILICAL VEIN;

RA L.I. , reader I.F. .
 RA "A novel serine protease from human umbilical vein endothelial
 RA cells " ;
 RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN
 RP SEQUENCE FROM N.A.
 RA Sheppard P. , Blumberg H. , Jelinek L. , Foster D. , O'Hara P. ;
 RA Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN
 RP SEQUENCE FROM N.A.
 RA TISSUE=UTERUS;
 RA Wambutt R. , Heubner D. , Mewes H.W. , Gassenhuber J. , Wiemann S. ;
 RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RL [4]
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=CERVIX CARCINOMA;
 RA Strausberg R. ;
 RC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RA -1 - SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY
 CC -1 - SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
 CC EMBL: AF015878 ; AAD01553.1 ; -.
 DR EMBL: AF193611 ; AAF07186.1 ; -.
 DR EMBL: AL136914 ; CAB66848.1 ; -.
 DR EMBL: BC001278 ; AAH01278.1 ; -.
 DR MPROB: S01 309. -.

DR	InterPro; IPR001314; Chymotrypsin.	RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.-H., Weitz C., Whittaker C., Wimling L., Wyrshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
DR	pfam; PF00089; trypsin; 1.	RA	"Functional annotation of a full-length mouse cDNA collection.";
SMART	PR00722; CHYMOTRYPSIN.	RA	RT
DR	SMART; SM00020; TYP_SPC; 1.	RA	Nature 409:685-690(2001).
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.	RL	-1 - SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
KW	Hydrolase; Protease; Serine protease.	CC	-1 - SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
SEQUENCE	383 AA; 43001 MW; 4666C11ABFD5E8F CRC64;	DR	EMBL; AK009847; BAB26541.1;
Query Match	100.0%; Score 2080; DB 4; Length 383;	DR	MGI; 1923703; 231006G15Rik.
Best Local Similarity	100.0%; Pred. No. 8_2e-14;	DR	InterPro; IPR001314; Chymotrypsin.
Matches	383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	DR	InterPro; IPR001254; trypsin.
Qy	1 MAGIPGLLFLFLFLCAVQVSPYSPAWKPTWPAVRLPVLPVLFQSTLNLAKEDEAAKLE 60	DR	pfaam; PF00089; trypsin; 1.
Db	1 MAGIPGLLFLFLFLCAVQVSPYSPAWKPTWPAVRLPVLPVLFQSTLNLAKEDEAAKLE 60	DR	PRINTS; PR00022; CHYMOTRYPSIN.
Qy	61 VSSSCGPQCHKGTPLPTEYEAKOYLSETLYANGSRPTEQYIYLSSSGDGAQHRDSGS 120	DR	SMART; SM00020; TRYSP; 1.
Db	61 VSSSCGPQCHKGTPLPTEYEAKOYLSETLYANGSRPTEQYIYLSSSGDGAQHRDSGS 120	DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
Qy	121 SGKSRKRKQIYGYDSSRFSTFGKDFDLYNPFSTPSVQLSTGCTGTLVAEKHVLTAACHTHDG 180	DR	KW
Db	121 SGKSRKRKQIYGYDSSRFSTFGKDFDLYNPFSTPSVQLSTGCTGTLVAEKHVLTAACHTHDG 180	DR	Hydrolase; Serine protease.
Qy	181 KTYVKGPKQLRQFLKPKFDGGRGANDSTSAMPQMKFQWLRVKRTHVPKCWKGNAND 240	DR	SEQUENCE 382 AA; 43053 MW; 69C9A7080E5B2306 CRC64;
Db	181 KTYVKGPKQLRQFLKPKFDGGRGANDSTSAMPQMKFQWLRVKRTHVPKCWKGNAND 240	DR	Query Match 90.6%; Score 1884.5; DB 11; Length 382;
Qy	241 IGMDYDALLEKKPKHFKEMIIGVSPAKQLPGRTRHSGYDNDREGNLVYRFDVKDE 300	DR	Best Local Similarity 90.6%; Pred. No. 9.2e-166; Matches 347; Conservative 15; Mismatches 20; Indels 1; Gaps 1;
Db	241 IGMDYDALLEKKPKHFKEMIIGVSPAKQLPGRTRHSGYDNDREGNLVYRFDVKDE 300	DR	Db 1 MAGIPGLLFLFLFLCAVQVSPYSPAWKPTWPAVRLPVLPVLFQSTLNLAKEDEAAKLE 60
Qy	301 TYDLYQOCDAQGASGSGVYRMWKRQOQKWERKLIGIFSGHOMGSPQDENAVVR 360	DR	Db 1 MAGIPGLLFLFLFLCAVQVSPYSPAWKPTWPAVRLPVLPVLFQSTLNLAKEDEAAKLE 60
Db	301 TYDLYQOCDAQGASGSGVYRMWKRQOQKWERKLIGIFSGHOMGSPQDENAVVR 360	DR	Db 1 VSSSCGPQCHKGTPLPTEYEAKOYLSETLYANGSRPTEQYIYLSSSGDGAQHRDSGS 120
Qy	361 ITPLYKAQICWYKGNYLDREG 383	DR	Qy 61 VSSSCGPQCHKGTPLPTEYEAKOYLSETLYANGSRPTEQYIYLSSSGDGAQHRDSGS 120
Db	361 ITPLYKAQICWYKGNYLDREG 383	DR	Db 60 VSSSCGPQCHKGTPLPTEYEAKOYLSETLYANGSRPTEQYIYLSSSGDGAQHRDSGS 120
RESULT	2	Qy	Db 120 TGRSRRKQIYGYDGRSIFGKDFLNNPFSTPSVQLSTGCTGTLVAEKHVLTAACHTHDG 179
Q9D6X6	PRELIMINARY; PRT; 382 AA.	Db	Qy 181 KTYVKGPKQLRQFLKPKFDGGRGANDSTSAMPQMKFQWLRVKRTHVPKCWKGNAND 240
AC	Q9D6X6; PRELIMINARY; PRT; 382 AA.	Db	Db 180 KTYVKGPKQLRQFLKPKFDGGRGANDSTSAMPQMKFQWLRVKRTHVPKCWKGNAND 239
DT	01-JUN-2001 (TREMBLrel. 17, Created)	Qy	Db 180 KTYVKGPKQLRQFLKPKFDGGRGANDSTSAMPQMKFQWLRVKRTHVPKCWKGNAND 239
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	Db	Qy 241 IGMDYDALLEKKPKHFKEMIIGVSPAKQLPGRTRHSGYDNDREGNLVYRFDVKDE 300
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	Db	Db 240 IGMDYDALLEKKPKHFKEMIIGVSPAKQLPGRTRHSGYDNDREGNLVYRFDVKDE 299
DE	2310046615RIK PROTEIN.	Qy	Qy 301 TYDLYQOCDAQGASGSGVYRMWKRQOQKWERKLIGIFSGHOMGSPQDENAVVR 360
GN	2310046615RIK.	Db	Db 300 TYDLYQOCDAQGASGSGVYRMWKRQOQKWERKLIGIFSGHOMGSPQDENAVVR 359
OS	Mus musculus (Mouse).	Qy	Qy 361 ITPLYKAQICWYKGNYLDREG 383
OC	Karyotype; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Db	Db 360 ITPLYKAQICWYKGNYLDREG 382
OX	NCBI_TaxID=1000;	RESULT	3
RN	[1]	Q9BQP6	PRELIMINARY; PRT; 413 AA.
RP	SEQUENCE FROM N.A.	ID	Q9BQP6
RC	STRAIN=C57BL/6J; TISSUE=TOONGUE;	AC	Q9BQP6;
RC	MEDLINE=21085660; PubMed=11217851;	DT	01-JUN-2001 (TREMBLrel. 17, Created)
RX	Kawai J., Shinagawa A., Shibusawa K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto I., Saito T., Okazaki Y., Gojobori T., Kasukawa T., Saito R., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Gissi C., King B., Kochiwa H., Kuehli P., Lewis S., Matsuo Y., Nikaido I., Pesse G., Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaize J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
RA	RA	DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
RA	RA	DE	DE223E3.1 (PUTATIVE SECRETED PROTEIN ZSIG13).
RA	RA	GN	DU223E3.1.
RA	RA	OS	OS Homo sapiens (Human).
RA	RA	OC	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RA	RA	OC	OC NCBI_TaxID=9606;
RA	RA	RN	RN [1]
RA	RA	RP	RP SEQUENCE FROM N.A.
RA	RA	RA	RA Dunn M.;
RA	RA	RL	RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RA	RA	DR	DR EMBL; ALI21939; CAC35071.1;
RA	RA	SQ	SEQUENCE 413 AA; 47098 MW; 818D9C951BD2D6C1 CRC64;

DR PROSITE; PS00672; V8_HIS; 1.
 DR PROSITE; PS00673; V8_SER; 1.
 KW Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 ?
 FT PROPEP 7 ? 94
 FT CHAIN 95 316 GLUTAMYL ENDOPEPTIDASE.
 FT ACT_SITE 141 141 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 261 261 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 126 142
 FT DISULFID 275 279
 SQ SEQUENCE 316 AA; 33611 MW; 96D7552CB7089B09 CRC64;

Query Match 9.1%; Score 194; DB 1; Length 316;
 Best Local Similarity 24.0%; Pred. No. 5e-09;
 Matches 79; Conservative 43; Mismatches 143; Indels 64; Gaps 15;

QY 63 SSCPQCHKGTPL --PYYEAKOYLSVETLYANGSRTETOQIYIILSSGGDAQHRSQG 119
 Db 28 AQAAQSPH -TPVYSDPSY-KAETSVTDP----NIKSDQYLYSKRAFTGIGKVNETKE 79

QY 120 SSGSRRRQYIYDSSRSITGDFELLN----YFETTSVPLVAEKHV 170
 Db 80 KAEEKSPAKAPY--SISVSVGDDDRTRVNTAYPRAVTHSSGSCWMMIGPKTV 136

QY 171 LTAACHTHDGKT-YVKGTRQKLRVGFLEPKFKDGGRGANDSTSAMPEQMKFQWTRVKRTHV 229

Db 137 ATAGHCYTDTSQSGSFAQPATVSPG-----RNGTS-----YPYGVKSTYFI 178

QY 230 PKGMIGKGNANDGMDYDALLELKKPHKRFKFMIGVSPPAKOLPGRTHFSGYDNDRPGN 289
 Db 179 PSSGRSGNTN----YDGAELSEPIGNTYGFYSYTTSSLVGTTVTSYGPDKTAG 233

QY 290 LYRFCD -VKDITYDLYQODCAQPGASGSYVWRKQROOKWERNII---GIFSGH 343
 Db 234 TQHQHSGPAAISETYKLOYAM-DTYGGOSGSPPYFEQSSSRTRNCSPGCSLAVHTNGVYGG- 291

QY 344 QWYDMNGSPQDFDNFVAVTRPLKQIQCYW 372
 Db 292 -----SSYNRGTRITKEVFDNLNW 311

RESULT 2
 MPR_BACSU
 ID MPR_BACSU STANDARD; PRT; 313 AA.
 AC P39790; 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE EXTRACELLULAR METALLOPROTEASE PRECURSOR (EC 3.4.21. -).
 GN MPR
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP STRAIN=G241;
 RX MEDLINE=90130256; PubMed=2105291;
 RA Sloma A., Rudolph C.F., Ruffo G.A. Jr., Sullivan B.J., Theriault K.A.,
 RA Ally D., Pero J.;
 RT "Gene encoding a novel extracellular metalloprotease in Bacillus
 subtilis";
 RT J. Bacteriol. 172:1024-1029(1990).
 RN [2] SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.,
 RA "Sequence analysis of the 70kb region between 17 and 23 degree of the
 Bacillus subtilis chromosome.";
 RT Submitted (S01-1997) to the EMBL/GenBank/DBJ databases.
 RL [3] SEQUENCE OF 1-68 FROM N.A.

RX MEDLINE=89108019; PubMed=3145906; Venema G.;
 RA Smith H., de Jong A., Bron S., Venema G.;
 RT "Characterization of signal-sequence-coding regions selected from the
 Bacillus subtilis chromosome.";
 RL Gene 70:351-361(1988).
 RN [4] SEQUENCE OF 1-10 FROM N.A.
 RP STRAIN=168;
 RX MEDLINE=96118702; PubMed=7496533;
 RA Saxild H.H., Jacobsen J.H., Nygaard P.;
 RT "Functional analysis of the *Bacillus subtilis* purt gene encoding
 formate-dependent glycynamide ribonucleotide transformylase.";
 RL Microbiology 141:2211-2218(1995).
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO PEPTIDE FAMILY S2B; ALSO KNOWN AS THE
 CC V8 FAMILY.
 CC -1- CAUTION: CALLED "METALLOPROTEASE", BUT CLEARLY BELONGS TO THE S2
 CC FAMILY OF SERINE PROTEASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC DR EMBL; I10505; AAA22604; 1; -.
 DR EMBL; AB006424; BAA33121; 1; -.
 DR EMBL; M22916; AAA22832; 1; -.
 DR EMBL; M299105; CAB12018; 1; -.
 PIR; A35122; A35122.
 DR HSPB; P00756; 1SGF.
 DR M6RPS; S01_272; -.
 DR SubList; B010690; mpfr.
 DR InterPro; IPR000126; Ser_proteas_V8.
 DR Pfam; PF00089; Lyspro; 1.
 DR PRINTS; PRO039; V8PROTEASE.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00672; V8_HIS; 1.
 DR PROSITE; PS00673; V8_SER; 1.
 KW Hydrolase; Serine protease; zymogen; Complete proteome.
 FT SIGNAL 1 34 AA; 33842 MW; D41788E8D652AE94 CRC64;
 FT PROPEP 35 93
 FT CHAIN 94 313 EXTRACELLULAR METALLOPROTEASE.
 FT ACT_SITE 146 146 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 267 267 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 131 131 BY SIMILARITY).
 FT CONFLICT 61 68. QVSAYEG -> PLESTAQA (IN REF. 3).
 SQ SEQUENCE 313 AA; 33842 MW; D41788E8D652AE94 CRC64;
 Query Match 6.4%; Score 135.5; DB 1; Length 313;
 Best Local Similarity 22.2%; Pred. No. 0.00038;
 Matches 77; Conservative 38; Mismatches 121; Indels 111; Gaps 17;
 QY 72 GIPPLPTYEAKOYLSVETLYANGSRTEQVGYIILSSSGD--GAQHRSGSSGKSRKRQ 129
 Db 29 GYPAKAENPQTQSVNTQGKEAATKNOT-----SKADQVSAPYRGTGKTSKS---- 75
 OY 130 TYGDSRF-----SIFGRD-----FLNYPFSTSVKLST-----GCTGTL 164
 Sq 313 AA; 33842 MW; D41788E8D652AE94 CRC64;
 Db 76 LGGQTEKEKNQTLQPSLIGIDERTISSFTSPPTVSVIPTNTSTYGTGFL 135
 QY 165 VAEKHLTAACIH-----DGKTYVKGTOKLVRGVFLKPKFQDGGGAND 208
 Db 136 VNPNTVYVAGCYSQDHGASTITAAPGRNGSSYYPT----- 175
 QY 209 STSAMPEMKFQWIRVRFTHVPGWIGNN-----ANDICMDYDVALLEKPKHFRKFMTIGV 265
 Db 176 -SGTMFYSVK-GWTESKDTNYDGAIKLNGSPGNTVCMW-YGRTNTNSSP-----VGL 225
 RP

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

! protein - protein search, using sw model

on : December 2, 2001, 16:17:55 ; search time 48.03 Seconds
 (without alignments)
 621.704 Million cell updates/s

title: US-09-072-384-15
 perfect score: 2131
 sequence: 1 MAGIPGLIPLLFFLICAVGQ.....IKGNYLDCREGDTVPPGSN 392

scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5.

searched: 219241 seqs, 76174552 residues

total number of hits satisfying chosen parameters: 219241

DB seq length: 0
 minimum DB seq length: 0
 maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

database : PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Length	DB ID	Description
1	194	9.1	A45134	endopeptidase (EC 3.4.24.12) metalloproteinase
2	135.5	6.4	A35122	coagulation factor
3	120.5	5.7	EXR5	pancreatic elastase
4	119	5.6	A26823	hypothetical protein
5	117	5.5	T29767	pancreatic elastase
6	115.5	5.4	B26823	pancreatic elastase
7	115	5.4	A25528	pancreatic elastase
8	115	5.4	S70439	pancreatic elastase
9	115	5.4	A5619	probable pancreatic elastase
10	114.5	5.4	T26615	trypsin-like proteinase
11	114	5.3	DBHU	complement factor
12	113	5.3	ELR1	pancreatic elastase
13	113	5.3	ELPG	pancreatic elastase
14	112.5	5.3	TRSMG	trypsin (EC 3.4.2.4)
15	112.5	5.3	E26823	hypothetical protein
16	112.5	5.3	H64915	hypothetical protein
17	111.5	5.2	EXHU	coagulation factor
18	111	5.2	JC5759	brain-specific serine protease
19	110	5.2	C26823	pancreatic elastase
20	110	5.2	EXBO	coagulation factor
21	110	5.2	T15308	hypothetical protein
22	109.5	5.1	T35117	probable secreted protein
23	108	5.1	A29154	complement factor
24	108	5.1	A47547	serine protease
25	108	5.1	A55617	T-cell suppressor
26	107.5	5.0	A28566	Ra-reactive factor
27	107.5	5.0	A59271	pancreatic elastase
28	105	4.9	S40054	trypsin related protein
29	105	4.9	ELRT2	pancreatic elastase

880	103.5	4.9	416	1	KFBO
881	102.5	4.8	1238	2	T34/29
882	100	4.7	272	2	JC4170
883	98.5	4.6	267	2	S400/6
884	98	4.6	548	2	D82/175
885	98	4.6	624	2	T302/89
886	97.5	4.6	409	2	T351/18
887	97	4.6	452	1	A30/51
888	97.5	4.6	747	2	I51/579
889	97	4.6	1019	2	A38/38
890	96.5	4.5	259	2	S68/24
891	96.5	4.5	275	2	I46/712
892	96.5	4.5	2145	2	JC4/47
893	95	4.5	430	1	A24/102
894	95	4.5	782	2	T32/155
895	94	4.4	268	2	S68/55

ALIGNMENTS

RESULT	1	A45134					
		C;Species: <i>Bacillus licheniformis</i>					
		C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999					
		C;Accession: A45134; S2378					
		C;Cross-references: GB:0060; NID:9216263; PIDN:BAA00949;1; PID:d1001415; PID:921626					
		C;Experimental source: ATCC 14580					
		C;Note: sequence extracted from NCBI backbone (NCBIN:118784; NCDBP:118785)					
		R;Svendsen, I.; Breddam, K.					
		Bur. J. Biochem. 204, 165-171, 1992					
		A;Title: Isolation and amino acid sequence of a glutamic acid specific endopeptidase					
		A;Reference number: S23078; MUID:92155199					
		A;Accession: S23078					
		A;Status: preliminary					
		A;Molecule type: protein					
		A;Residues: 95-316 <SVE>					
		C;Keywords: hydrolase					
		Query Match 9.1%; Score 194; DB 2; Length 316;					
		Best Local Similarity 24.0%; Fred. No. 1.4e-08;					
		Matches 79; Conservative 43; Mismatches 143; Indels 64; Gaps 15;					
Qy	63	SSCGPQCHKCTPL-PTVEBAKOLSYESTLYANGSRTETQVGYIILSSSGDGAQHRSQG 119					
Db	28	AQAAABSPH--TPVSSSDPSY KAENSTVYDP-----NIKSDQYGLYSSKAFTGTGKVNETKE 79					
Qy	120	SSGKSRRKQIYGYDSRFISFGKDPLLNN-----YFSTSVKLST---GCTGTLYAEKHY 170					
Db	80	KAEEKSPAKAPY--SIKSVIGSDDRTTRYNTNTAYPYRAVHSSISGCTGMMIGPKIV 136					
Qy	171	TAHHETHDKST-YWKGTQKLRYGFLKPKFKDGGHAGNDSTSAPEQMKFOWTRVKRTHV 229					
Db	137	ATAGCICYDTS-SSPAGTATVSPG-----RNGTS-----YYGSVSRSTRFY 178					
Qy	230	PKGWTKGKGNANDIGMDYDALLELKKPHKRMKFMIGVSPPAKQLPQGGRHFGSGYDNDRPGN 289					
Db	179	PSGNFSGNTN-----YDGAIELSEPIGNTYVGFGYSYTSSVLNGTTVSGPGDKTAG 233					
Qy	290	LYVYRCFD--TKDENTDLYQQCDAGPGASGSGYVYRMVKRQQQKWERKII---GIFSGH 343					
Db	234	TQWQHSGPIA SEYVQLQIAM-DYGGQSGSPVPOQSSRTINCSPCSLAVHTNGVYGA- 291					

F:103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
 F:187/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:208/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231-332/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #status predicted
 F:274-320,417/Active site: His, Asp, Ser #status predicted

Query Match 5.71; Score 120.5; DB 1; Length 482;
 Best Local Similarity 32.3; Pred. No. 0.03; Indels 23; Gaps 4;
 Matches 31; Conservative 17; Mismatches 23; Indels 25; Gaps 4;
 C:Genetics: A:Gene: CESP-ZC581.6
 Db 160 CGTGLVAEKKVTTAAACIHDTGKTYVKGTLRKPFDGGRGANDSISAMPEQMKF 219
 Qy 220 QMIRVKRTHVYKTTAAACIHDTGKTYVKGTLRKPFDGGRGANDSISAMPEQMKF 219
 Db 259 CGGTILNEFYVTTAAACIHDTGKTYVKGTLRKPFDGGRGANDSISAMPEQMKF 219
 Db 313 Q----RDTY-----DPIIAMLRLKTP 329

RESULT 4

A26823

pancreatic elastase II (EC 3.4.21.71) precursor - pig

C:Species: Sus scrofa domestica (domestic Pig)

C:Date: 16-Aug-1988 #sequence_change 22-Jun-1999

C:Accession: A26823

R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.

DNA 6, 163-172, 1987

A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed

A:Reference number: A90958; MUID:87217962

A:Accession: A26823

A:Residues: 1-269 <KAW>

A:Cross-references: GB:MI16651; NID:916441; PID:916442

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-28/Domain: propeptide #status predicted <PRO>

F:29-289/Product: elastase II #status predicted <MAT>

F:29-362/Domain: trypsin homology <TRY>

F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 5.68; Score 119; DB 2; Length 269;

Best Local Similarity 27.5%; Pred. No. 0.02; Indels 66; Gaps 14;

Matches 60; Conservative 28; Mismatches 64; Indels 66; Gaps 14;

Db 128 RQIYGDSRFISFGKDFLNNPFSTSVKL-STG-----CTGTVAEHHVLTAAHCTDGG 181

Db 28 RIVGGEDARPN -----SFWQVSLOYDSSGQWHRCTGTLVQDSSWVLTAAHCTSSR 79

Db 182 TY-VKGTKVLRGVFLKPKFDDGGRGANDSTSAMPQMKFQWIRVKRTHVKGWIKGNAN 239

Db 80 TYRVVLRGRHSL-----STNEPGLA-----VTKVSKLVAHQDW----NSN 114

Db 240 DGMDDYDALLKLPK HKRKFMKIGVSPPKQ-LPG-----GRIHHSQYDND--R 286

Db 115 QLSNGNDIALKLKASPLS LTDKIOLGCLPAAGTILEPNNYCYXTGNGRLQTNGASFDILQ 174

Db 287 PGNLVYRFCDYKDETDLYQOCDAQFGASGSQYVYRM 324

Db 175 QGQLL-----YATC-SKPGWNGSTVKNM 200

Db 358 PHEIYFQARDITDTKTVACVVLKILNKTOASLNLSKGDSGGAAIDV-----KGK 411

Db 334 RKLIGITS-----GHWQDMNGSPQDFNVAVRTPLYAQIC 370

Db 412 KTIIGLVSQTSQKRRQSGNETMELSYGVFTKQI-----CKYTGIC 453

RESULT 6

B26823 pancreatic elastase II (EC 3.4.21.71) A precursor - human

C:Species: Homo sapiens (man)

C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 20-Jun-2000

C:Accession: A26823; A27432; A1431; S34491

R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.

DNA 6, 163-172, 1987

A:Title: Characterization of pancreatic elastase 2 determined by sequence analysis

A:Reference number: A90958; MUID:87217962

A:Accession: B26823

A:Molecule type: mRNA

A:Residues: 1-269 <KAW>

A:Cross-references: GB:MI16652; NID:9182057; PID:AAA52380.1; PID:9182058

R:Fletcher, T.S.; Shen, W.P.; Largman, C.

Biochemistry 26, 7256-7261, 1987

A:Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis

A:Accession: A27432

A:Molecule type: mRNA

A:Residues: 1-269 <FILE>

RESULT 5

T29767 hypothetical protein ZC581.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29767

R:Waterston, R.; Gattung, S.; Le, T.T.

A;Cross-references: GB:MI6631; NID:g182022; PIDN:AAA52374_1; PID:g182023
 R; Shirasu, Y.; Yoshida, H.; Matsuki, S.; Takemura, K.; Ikeda, N.; Shimada, Y.; Ozawa, T.
 J. Biochem. 102, 1555-1563, 1987

A;Title: Molecular cloning and expression in Escherichia coli of a cDNA encoding human F
 A;Reference number: A41431; MUID:88198076

A;Accession: A41431
 A;Molecule type: mRNA
 A;Residues: 1-201, V, 203-269 <SH1>
 A;Cross-references: GB:D00236; NID:g219619; PIDN:BA00165_1; PID:g219620
 A;Note: the authors translated the codon GAG for residue 202 as Cys
 R; Moullé, M.; Michon, T.; Kerfèlec, B.; Chapus, C.
 FEBS Lett. 261, 179-183, 1990

A;Title: Further studies on the human pancreatic binary complexes involving procarboxypeptidase
 A;Reference number: S08253; MUID:9016911
 A;Cross-references: GDB:ELA1
 A;Molecule type: protein
 A;Residues: X, 18-50 <NOU>
 C;Genetics:
 A;Gene: GDB:ELA1
 A;Cross-references: GDB:119866; OMIM:130120
 A;Map position: 12pter-12qter
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; serine proteinase
 F;1-16/Domain: signal; signal sequence #status predicted <SIG>
 F;17-28/Domain: propeptide #status predicted <PRO>
 F;29-269/Domain: pancreatic elastase II A #status predicted <MAT>
 F;29-262/Domain: trypsin homology <TRY>
 F;73,121-216/Active site: His, Asp, Ser #status predicted

Query Match 5.4% Score 115.5; DB 2; Length 269;
 Best Local Similarity 28.6%; Pred. No. 0.038; Gaps 11;
 Matches 50; Conservative 21; Mismatches 51; Indels 53; Gaps 11;

QY 147 NYPESTSYKLSTG-----CTGTIVAEFHVLAAHCIDHGKTYVKGTOKLRYGFLKPKF 200
 DB 39 SWPQVQSVLSOYSSNGKWHHTCGGLSIANSVLAHCISSSR-----RVL----- 85

QY 201 DGGRG---ANDSTSAMPEQMKFQWIRVTPKGWIKGNANDIGMDYDYLALLELKPK- 255
 DB 86 -GRHNLYVAEGLSA-----VSYSKIVYHKDW---NSNQISKGNDIALLKLANPV 131

QY 256 HKRKEMKIGVSPPAKQ-LPG-----GRIHESGYND-RPGNLY--YRFC 295
 DB 132 SLIDKIQLACPPATLDPNNYPCYVGMGRQLTNGAVPDVLOGRLLVVDATC 186

RESULT 7
 A;Cross-references: GB:X04573; NID:g50825; PIDN:CAA28242_1; PID:g50826
 C;Species: Mus musculus (house mouse)
 C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
 C;Accession: A25528
 R; Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
 R; Nucleic Acids Res. 14, 8307-8330, 1986
 A;Title: Sequence organisation and transcriptional regulation of the mouse elastase II
 A;Reference number: A93646; MUID:87066713
 A;Accession: A25528
 A;Molecule type: mRNA
 A;Residues: 1-271 <STE>
 A;Cross-references: GB:X04573; NID:g50825; PIDN:CAA28242_1; PID:g50826
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; serine proteinase
 F;1-10/Domain: signal sequence #status predicted <SIG>
 F;31-271/Product: pancreatic elastase II #status predicted <MAT>
 F;31-264/Domain: trypsin homology <TRY>
 F;75,123,218/Active site: His, Asp, Ser #status predicted

Query Match 5.4% Score 115.5; DB 2; Length 271;
 Best Local Similarity 25.0%; Pred. No. 0.039; Gaps 18; Mismatches 38; Indels 83;

QY 148 YPFSTSVK-LSTG-----CTGTIVAEFHVLAAHCIDHGKTY-VKGTOKLRVGFLLPKF 199
 DB 42 WPNOVSVQYQVLSGGRHNGGSIVANNNVLTAAHCLRYQTYRVLGHSL----- 92

QY 200 KDGGRGANDSTSAMPEQMKFQWIRVTPKGWIKGNANDIGMDYDYLALLELKPK- 258
 DB 93 -SNPGAGSAA-----VQVSKLVHQHN--NSQNVNGYDYLALIKLASPVTLIS 136

QY 259 KEMKIGVSPPAKQ-LPGGRI-HFSGY-----YRFC----- 295
 DB 137 KNIQACLPPAGTLLPRNIVCYVIGWGLLQTANGNSPDTLROGRLLVVDATC 196

QY 296 DVKDETY---DLYQCDQDAQPGASGSVYVVRMVKRQOQKWERKIGFSGHQWDMNGS 351
 DB 197 SVKSSMVCAAGGDGTYSSCN---GDSGPINCRASNGO--WQ--VHGTVSFGLGCN-Y 247

QY 352 PQDENAVVIRTPKYAQICYWKGNYL 379
 DB 248 PRKESVFTRVS-----NYID 262

RESULT 8
 S70439
 pancreatic elastase I (allele HEL-16) probable splice form I - human
 C;Species: Homo sapiens (man)
 C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 16-Dec-1998
 C;Accession: S70439
 R; Kawashima, I.; Tani, T.; Shimoda-Takano, K.; Ohmine, T.; Furukawa, DNA Seq. 2, 303-312, 1992
 A;Title: Genomic organization of the human homologue of the rat pancreatic elastase I
 A;Reference number: A56615; MUID:9233395
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-258 <KAW>
 C;Superfamily: trypsin; trypsin homology <TRY>
 F;19-251/Domain: trypsin homology <TRY>

Query Match 5.4%; Score 115; DB 2; Length 258;
 Best Local Similarity 26.9%; Pred. No. 0.04; Gaps 6; Mismatches 35; Conservative 23; Indels 34; Gaps 6;

QY 147 NYPESTSYKLSTG-----CTGTIVAEFHVLAAHCIDHGKTY-VKGTOKLRVGFLLPKF 198
 DB 29 SWPQISLQYRSGGSWYITCGGTILRQWVMTAAHCVYQKTPRVAVDHNL----- 80

QY 199 FKDGGRGANDSTSAMPEQMKFQWIRVTPKGWIKGNANDIGMDYDYLALLEL-KKPHK 257
 DB 81 -----SQNDGTE-----QVYSVQKIVVHPW--NSDNVAAQYDIALRLAOSTV 123

QY 258 RKEMKIGVSPPAKQ-LPG-----GRIHESGYND-RPGNLY--YRFC 267
 DB 124 NSTVQLGVLP 133

RESULT 9
 A;Accession: A56615
 probable pancreatic elastase (EC 3.4.21.36) pseudogene - human
 N;Alternate names: pancreatic elastase I homolog; pancreatic elastase I allele HEL1-1
 C;Species: Homo sapiens (man)
 C;Date: 11-Aug-1995 #sequence_revision 17-Aug-1995 #text_change 14-Aug-1998
 C;Accession: A56615; S70440
 R; Kawashima, I.; Tani, T.; Mita-Honjo, K.; Shimoda-Takano, K.; Ohmine, T.; Furukawa, DNA Seq. 2, 303-312, 1992
 A;Title: Genomic organization of the human homologue of the rat pancreatic elastase I
 A;Reference number: A56615; MUID:9233395
 A;Accession: A56615
 A;Molecule type: DNA
 A;Residues: 1-267 <KAW>
 A;Cross-references: EMBL:X62259; NID:431246; EMBL:X62258; GB:SA0923; NID:931247; EMBL:5; GB:SA0856; NID:931251; EMBL:X62256; GB:SA0857; NID:93152; EMBL:62257; GB:SA0859;
 A;Note: sequence extracted from NCBI backbone (NCBIN:109315, NCBIN:109317, NCBIN:1093

C;Comment: This apparently silent human homolog of pancreatic elastase I is a single-copy protein in some other tissue.

C;Genetics:

A;Gene: GDB:ELA1

A;Map Position: 12

C;Keywords: hydrolase; pseudogene; serine proteinase

Query Match 5.4%; Score 115; DB 4; Length 267;
Best Local Similarity 26.9%; Pred. No. 0.042; Mismatches 38; Indels 34; Gaps 6;

Db 38 SWPSQISLQYRGGSWHTCGTLIRQWVTAACVQYDQTFRVVAGDHLN----- 89

Qy 199 FKDGGRGANDSTSAMPEQMKFQWIRVKRTHPKGWIKGNNANDGMDYDALLEL-KKPHK 257

Db 90 -----SQDCTE-----QYVSQKIVVHPW-----NSDVAAGCYDIALURLAQSVTL 132

Qy 258 RKFMKIGVSP 267

Db 133 NSYVQLGVLP 142

RESULT 10
TRW5Y

trypsin-like proteinase (BC 3.4.21.-) 5G1 precursor - yellow fever mosquito (fragment)

C;Species: Aedes aegypti (yellow fever mosquito)

C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999

A;Accession: S19891

R;Khalokh, S. ; Tabak, L.M. ; Prosser, D.E. ; Downe, A.E.R. ; White, B.N.
submitted to the EMBL Data Library, February 1992

A;Description: Isolation, sequencing and characterization of 2 cDNA clones coding for trypsin-like proteinase (BC 3.4.21.-) 5G1 precursor - yellow fever mosquito (fragment)

A;Accession: S19890

A;Molecule type: mRNA

A;Residues: 1-238 <HLA>

A;Cross-references: EMBL:X64363; NID:g5563; PID:CAA45715.1; PID:g5564

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; insect midgut; protein digestion; serine proteinase

F:1-11/Domain: signal sequence and propeptide (fragment) #status predicted

F:12-232/Domain: trypsin homology <TRY>

F:38-54,162-178,189-213/disulfide bonds; #status predicted

F:53,97,193/Active site: His, Asp, Ser #status predicted

Query Match 5.4%; Score 114.5; DB 1; Length 238;
Best Local Similarity 29.4%; Pred. No. 0.04; Mismatches 44; Indels 35; Gaps 7;

Db 24 PQVSLSGVGSSHFCCGSLISERWVNTAGHQAASGQNL---QVHIG---SSQHASSG-- 75

Qy 206 ANDSTSAMPEQMKFQWIRVKRTHPKGWIKGNNANDGMDYDALLELKKPHKRRKEMKIG 264

Db 76 -----QIYKVKVNRHPK-----YDEVTTDYLDFALLEEFTVTFSDSCAP 115

Qy 265 VSPPAKQLP---GGRTHFSGKD 284

Db 116 VKLPOKDTPVNEGTCLOVSGGN 138

RESULT 12

ELRT1

pancreatic elastase (EC 3.4.21.36) I precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 18-Jun-1999

C;Accession: A0060; A20534

R;MacDonald, R.J. ; Swift, G.H. ; Quinto, C. ; Swain, W. ; Pictet, R.L. ; Nikovits, W. ; Ru

Biochemistry 21, 1453-1463, 1982

A;Title: Primary structure of two distinct rat pancreatic proelastases determined

A;Reference number: A00960; MUID:82182967

A;Accession: A0060

A;Molecule type: mRNA

J. Biol. Chem. 267, 9210-9213, 1992
A;Title: Human adipins is identical to complement factor D and is expressed at high 1

A;Reference number: A40197; MUID:92220520

A;Accession: A40197

A;Molecule type: mRNA

A;Residues: 1-246 <W11>

A;Cross-references: GB:MB84526

R;Niemann, M.A. ; Bhown, A.S. ; Bennett, J.C. ; Volanakis, J.E.

Biochemistry 23, 2482-2486, 1984

A;Title: Amino acid sequence of human D of the alternative complement pathway.

A;Reference number: A00936; MUID:85000441

A;Accession: A00936

A;Molecule type: protein

A;Residues: 19-44, 'G', '46-51, 'Q', '53-75, 'P', '80-83, 'XXXIT', '90-172, '86-91, '185-23

A;Note: a few residues were assigned from the previously published sequence of Reid et

R.Miyata, T. ; Oda, O. ; Inagi, R. ; Sugiyama, S. ; Miyama, A. ; Maeda, K. ; Nakashima, I. ; Mol. Immunol. 27, 637-644, 1990

A;Title: Molecular and functional identification and purification of complement compo

A;Reference number: A60571; MUID:90370044

A;Accession: A60571

A;Molecule type: protein

A;Residues: 19-20, 'XX', '23-27, 'XX', '30-31, 'XX', '34, 'X', '36-40 <MY>

R;Balke, N. ; Holtkamp, U. ; Hoerl, W.H. ; Tschesche, H.

FEBS Lett. 371, 300-302, 1995

A;Title: Inhibition of degranulation of human polymorphonuclear leukocytes by comple

A;Reference number: S66645; MUID:96013156

A;Accession: S66645

A;Status: preliminary

A;Molecule type: protein

A;Residues: 19-44, 'C', '46-48 <BAL>

C;Comment: Factor D cleaves factor B when the latter is complexed with factor C3b, ac

C;Genetics:

A;Gene: GDB:DF

A;Cross-references: GDB:132645; OMIM:134350

A;Map position: Xpter-Xqter

C;Superfamily: trypsin homology

C;Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-246/Product: complement factor D (fragment) #status experimental <MAT>

F:19-21/Domain: trypsin homology <TRY>

F:44-60,141-207,172-188,197-222/Disulfide bonds: #status predicted

F:59,105,201/Active site: His, Asp, Ser #status predicted

Query Match 5.3%; Score 114; DB 1; Length 246;
Best Local Similarity 28.4%; Pred. No. 0.046; Mismatches 43; Indels 48; Gaps 9;

Qy 149 PESTSYKLSTG---CTGTLVAEKKVHLTAAHCIDHGRKYVKTQKLRYGLPKFKDGG 203

Db 31 PYMASVQLQNLGAHLCAGWVLAERWVLSAHCLEDAADGKVQL----- 72

Qy 204 RGANDSTSAMPEQMKFQWIRVKRTHPKGWIKGNNANDGMDYDALLELKKPHKRRKEMKIG 263

Db 73 LGAHSLSQPEPSKRLYLVRA---VP---HPDSQDTIDHLILQSE-----KA 117

Qy 264 GVSPPAKQLPGRTHFSGYDND-RPGNLVYRFCDV 297

Db 118 TLGPAVRPLPQRV---DRDVAPGTL---CDV 143

RESULT 12

DBHU

complement factor D (EC 3.4.21.46) precursor [validated] - human (fragment)

C;Alternate names: adipins (man); C3 convertase activator

C;Species: Homo sapiens (man)

C;Date: 28-Aug-1985 #sequence_revision 31-Dec-1992 #text_change 08-Dec-2000

C;Accession: A40197; A00936; A60571; S66645

R;White, R.T. ; Damm, D. ; Hancock, N. ; Rosen, B.S. ; Lowell, B.B. ; Usher, P. ; Flier, J.S.

F;17-26/Product: activation peptide #status predicted <APT>	Query Match	5.3%	Score 113; DB 1; Length 27
F;21-259/Product: elastase I #status experimental <MAT>	Best Local Similarity 27.3%; Pred. No. 0.061; Mismatches 41; Indels 36; Conservative 21; Mismatches 41; #status experimental		
F;27-259/Domain: trypsin homology <TRY>			
F;56-72,153-220,184-200,210-240/Disulfide bonds: #status experimental			
F;71,119,214/Active site: His, Asp, Ser #status experimental			
Qy 147 NYPFSTSYKLSTG 5.3% Score 113; DB 1; Length 27	Qy 147 NYPFSTSYKLSTG 5.3% Score 113; DB 1; Length 27		
Best Local Similarity 27.3%; Pred. No. 0.061; Mismatches 41; Indels 36; Conservative 21; Mismatches 41; #status experimental	Best Local Similarity 27.3%; Pred. No. 0.061; Mismatches 41; Indels 36; Conservative 21; Mismatches 41; #status experimental		
Db 37 SWSQISLQYRSGSSWAHTCGGTILRQNWYMTAAHCYDRELTFRVYVG 5.3% Score 113; DB 1; Length 27	Db 37 SWSQISLQYRSGSSWAHTCGGTILRQNWYMTAAHCYDRELTFRVYVG 5.3% Score 113; DB 1; Length 27		
Qy 199 FKGDRGRGANDSTSAMPENQKQFWIRVKRHRVPKGWIKGNANDIGMDYD 5.3% Score 113; DB 1; Length 27	Qy 199 FKGDRGRGANDSTSAMPENQKQFWIRVKRHRVPKGWIKGNANDIGMDYD 5.3% Score 113; DB 1; Length 27		
Db 89 -----NQDGTE -----QIVGVQKRVVHPW--NTDDVVAQY 5.3% Score 113; DB 1; Length 27	Db 89 -----NQDGTE -----QIVGVQKRVVHPW--NTDDVVAQY 5.3% Score 113; DB 1; Length 27		
Qy 258 RKFMKIGYSPPA 269 5.3% Score 113; DB 1; Length 27	Qy 258 RKFMKIGYSPPA 269 5.3% Score 113; DB 1; Length 27		
Db 132 NSVQLGVLPRA 143 5.3% Score 113; DB 1; Length 27	Db 132 NSVQLGVLPRA 143 5.3% Score 113; DB 1; Length 27		
RESULT 14	RESULT 14		
TRMG	trypsin (EC 3.4.21.4) precursor - Streptomyces griseus		
C:Species: Streptomyces griseus	C:Species: Streptomyces griseus		
C:Date: 24-Apr-1984 #sequence_revision 12-May-1994 #text_ch	C:Date: 24-Apr-1984 #sequence_revision 12-May-1994 #text_ch		
C:Accession: JQ1302; A00962	C:Accession: JQ1302; A00962		
R:Name: J.C.; Cha, S.H.; Jeong, S.T.; Oh, S.K.; Byun, S.M.	R:Name: J.C.; Cha, S.H.; Jeong, S.T.; Oh, S.K.; Byun, S.M.		
Biochem. Biophys. Res. Commun. 181, 707-713, 1991	Biochem. Biophys. Res. Commun. 181, 707-713, 1991		
A:Title: Molecular cloning and nucleotide sequence of Streptomyces griseus trypsin	A:Title: Molecular cloning and nucleotide sequence of Streptomyces griseus trypsin		
A:Reference number: JQ1302; MUID:92095977	A:Reference number: JQ1302; MUID:92095977		
A:Accession: JQ1302	A:Accession: JQ1302		
A: Molecule type: DNA	A: Molecule type: DNA		
A:Residues: 1-259 <KIM>	A:Residues: 1-259 <KIM>		
A:Cross references: GB M64471	A:Cross references: GB M64471		
A:Experimental source: strain ATCC10137	A:Experimental source: strain ATCC10137		
R:Olfafson, R.W.; Jurasek, L.; Carpenter, M.R.; Smillie, L. E.	R:Olfafson, R.W.; Jurasek, L.; Carpenter, M.R.; Smillie, L. E.		
Biochemistry 14, 1168-1177, 1975	Biochemistry 14, 1168-1177, 1975		
A:Title: Amino acid sequence of Streptomyces griseus trypsin	A:Title: Amino acid sequence of Streptomyces griseus trypsin		
A:Reference number: A00962; MUID:75127940	A:Reference number: A00962; MUID:75127940		
A: Molecule type: Protein	A: Molecule type: Protein		
A:Residues: 37-95,98-259 <OLA>	A:Residues: 37-95,98-259 <OLA>		
R: Read, R.J.; James, M. N. G.	R: Read, R.J.; James, M. N. G.		
J. Mol. Biol. 200, 523, 1988	J. Mol. Biol. 200, 523, 1988		
A:Title: Refined crystal structure of Streptomyces griseus	A:Title: Refined crystal structure of Streptomyces griseus		
A:Reference number: A44574; MUID:88286735	A:Reference number: A44574; MUID:88286735		
A:Contents: annotation; X-ray crystallography, 1.7 angstrom	A:Contents: annotation; X-ray crystallography, 1.7 angstrom		
A:Note: residues 96-97 modeled to Gly-Ala	A:Note: residues 96-97 modeled to Gly-Ala		
C:Genetics:	C:Genetics:		
A:Gene: sprt	A:Gene: sprt		
C:Superfamily: trypsin; trypsin homology	C:Superfamily: trypsin; trypsin homology		
C:Keywords: hydrolase; serine proteinase	C:Keywords: hydrolase; serine proteinase		
F;1-32/Domain: signal sequence #status predicted <SIG>	F;1-32/Domain: signal sequence #status predicted <SIG>		
F;33-36/Domain: propeptide #status predicted <PRO>	F;33-36/Domain: propeptide #status predicted <PRO>		
F;37-258/Product: trypsin #status experimental <MAT>	F;37-258/Product: trypsin #status experimental <MAT>		
F;73-254/Domain: trypsin homology <TRY>	F;73-254/Domain: trypsin homology <TRY>		
F;58-74,177-192,204-235/Disulfide bonds: #status experimental	F;58-74,177-192,204-235/Disulfide bonds: #status experimental		
F;73,118,208/Active site: His, Asp, Ser #status experimental	F;73,118,208/Active site: His, Asp, Ser #status experimental		
Qy 148 YPFSTSYKLSTGCTGTLVAKHVLTAACIHDGKTYVKGTQKLVRVGFELK 5.3% Score 112.5; DB 1; Length 26	Qy 148 YPFSTSYKLSTGCTGTLVAKHVLTAACIHDGKTYVKGTQKLVRVGFELK 5.3% Score 112.5; DB 1; Length 26		
Best Local Similarity 26.0%; Pred. No. 0.065; Mismatches 51; Indels 39; Conservative 21; Mismatches 53; #status experimental	Best Local Similarity 26.0%; Pred. No. 0.065; Mismatches 51; Indels 39; Conservative 21; Mismatches 53; #status experimental		
Db 48 FPF- -MYSMGGGALYAQDVLTAACV 5.3% Score 112.5; DB 1; Length 26	Db 48 FPF- -MYSMGGGALYAQDVLTAACV 5.3% Score 112.5; DB 1; Length 26		
Qy 208 DS---TSAMPENQKQFWIRVKRHRVPKGWIKGNANDIGMDYDALLEK 5.3% Score 112.5; DB 1; Length 26	Qy 208 DS---TSAMPENQKQFWIRVKRHRVPKGWIKGNANDIGMDYDALLEK 5.3% Score 112.5; DB 1; Length 26		

RESULT 15

Db 82 TSTATGGVVDLQSSAVKVRSTKVLQA---PGYNGTGDWALIKLAQPINQPTLKIA 136

Qy 265 WSPPAKQLPGGRHFSQYDNRPGNLYVRF 294

Db 137 TTYAYNO---GTETVAGWANREGGSQRY 163

Db 85765 hypothetical protein Z2592 [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C:Accession: E85765

R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, R.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7

A:Reference number: A85480; PMID:21074935; PMID:11206551

A:Accession: E85765

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-273 <STOP>

A:Cross-references: GB:AE005174; NID:912515576; PIDN:AA56585.1; GSPDB:GN00145; UWGP:Z25

A:Experimental source: strain O157:H7, substrate EDL933

C:Genetics:

A:Gene: Z2592

Query	Match	Score	Length
Best Local	Similarity	23.3%	273
Matches	Conservative	35	Pred. No. 0.069;
		5	Mismatches 112;
		57	Indels 57;
		13	Gaps 13;
Qy	101 VGYIILSSSGSSKSQRHRSGSQSSKSRRRQIIGYDSFSTFGDFLLNYPFESTSVKUSTG-	5.38	112.5;
Db	9 LGAISLTSASFVADPKDVAKSANDEVSTLFGHDDRPV--NDTQSPWMAVGQLETA	23.3%	2
Qy	160 ---CTGTLYAEEKHVLTAACHTDGGKTYVKGQKLRVGFLLKPKFKDGR-----GANDS	209	273;
Db	66 GNCIATLAPIALTAGHCL---LTPPKGKADKAYALRFVSNKGWRYDHIDIEGRVDP	159	
Qy	210 TSAMPEOMKFWQIVRKRTVHPKGWIGKNNANDIGMDYALLELKPKHKMKIGVSP--	23.3%	
Db	123 T---LGKRLKAD-----GDGWIYPPA---AAPNDFGLIVLRNFP-----GTPLP	267	
Qy	268 -----PAKQLPGRGIIFSGYDIDRPGNLY-RCFDYLDLQQCDQPGA	209	
Db	163 LFEQDKAALTAAALKAAAGRKVHQAGYPPDHDTLYSHONCEVTGWAQTSVMSHQCDTLPGD	242	315
Qy	316 SSGGYYVRMWKRQQQKWERKLIGIFS	341	
Db	223 SGSPML---HTDDGWM---LIGVQS	242	222

Search completed: December 2, 2001, 16:17:56
 Job time: 191 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:14:40 ; Search time 35.78 Seconds
(without alignments)

246.542 Million cell updates/sec

Title: US-09-072-384-2
Perfect score: 2112
Sequence: 1 MAGIPGSLFLFLCAVGQ IKGNYLDCREGDTVFLPGSN 392

Scoring table: BLOSUM62
Gapext 10.0 , Gapext 0.5

Searched: 212252 seqs, 2503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:
1: /cgn2_6/pcdatadata/2/1aa/5A_COMB.pep:
2: /cgn2_6/pcdatadata/2/1aa/5B_COMB.pep:
3: /cgn2_6/pcdatadata/2/1aa/6A_COMB.pep:
4: /cgn2_6/pcdatadata/2/1aa/6B_COMB.pep:
5: /cgn2_6/pcdatadata/2/1aa/PCTUS_COMB.pep:
6: /cgn2_6/pcdatadata/2/1aa/backfile1.pep:
*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2104	99.6	392	4	US-09-072-384-2	Sequence 2, Appli
2	2085	98.7	392	4	US-09-072-384-15	Sequence 15, Appli
3	2044	96.8	383	4	US-09-072-384-18	Sequence 1, Appli
4	165	7.8	222	1	US-08-090-048-1	Sequence 1, Appli
5	165	7.8	222	2	US-08-092-550-1	Sequence 1, Appli
6	165	7.8	222	2	US-07-927-661A-1	Sequence 1, Appli
7	123	5.8	256	3	US-08-066-769-89	Sequence 89, Appli
8	123	5.8	256	3	US-08-066-616-89	Sequence 89, Appli
9	123	5.8	256	4	US-08-017-795-89	Sequence 89, Appli
10	123	5.8	256	4	US-08-039-075A-89	Sequence 89, Appli
11	123	5.8	256	4	US-09-012-431-89	Sequence 89, Appli
12	123	5.8	256	4	US-09-032-215-32	Sequence 32, Appli
13	123	5.8	256	4	US-09-012-612-89	Sequence 89, Appli
14	114	5.8	256	4	US-08-066-613-89	Sequence 89, Appli
15	123	5.8	256	5	PCT-US5-14442A-89	Sequence 2, Appli
16	121.5	5.8	437	1	US-08-087-017-2	Sequence 59, Appli
17	1118	5.6	241	4	US-08-044-483-59	Sequence 3, Appli
18	116.5	5.5	437	1	US-08-087-037-3	Sequence 3, Appli
19	114	5.4	254	1	US-08-330-978-3	Sequence 3, Appli
20	114	5.4	254	1	US-08-074-042-3	Sequence 3, Appli
21	114	5.4	254	1	US-08-084-538-3	Sequence 3, Appli
22	114	5.4	254	1	US-08-074-532-3	Sequence 3, Appli
23	114	5.4	306	1	US-08-030-978-1	Sequence 1, Appli
24	114	5.4	306	1	US-08-434-042-1	Sequence 1, Appli
25	114	5.4	306	1	US-08-084-538-1	Sequence 1, Appli
26	114	5.4	306	1	US-08-074-592-1	Sequence 1, Appli
27	114	5.4	448	1	US-08-295-411-3	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-072-384-2
; Sequence 2, Application US/09072384
; Patent No. 6153420

GENERAL INFORMATION:
; APPLICANT: Shappard, Paul O.
; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
; TITLE OF INVENTION: AND MATERIALS AND METHODS FOR MAKING THEM
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
; CURRENT APPLICATION NUMBER: US/09/072, 384
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E.
; REGISTRATION NUMBER: 31, 648
; REFERENCE/DOCKET NUMBER: 91-16C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1..19
; OTHER INFORMATION:
; US-09-072-384-2

Query Match 99.6%; Score 2104; DB 4; Length 392;
 Best Local Similarity 100.0%; Pred. No. 6.6e-219; ;
 Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0; ;
 Feature: ;
 NAME/KEY: Signal Sequence ;
 LOCATION: 1..19 ;
 OTHER INFORMATION: ;
 US-09-072-384-15

QY 1 MAGIGLGLFLFLFLICAVGQVSPYSAWPWKTWPAVRLPVVLPPOSTNLAKPDKFGAEAKLE 60
 1 MAGIGLGLFLFLFLICAVGQVSPYSAWPWKTWPAVRLPVVLPQSTNLAKPDKFGAEAKLE 60
 Db 61 VSSSGPQCHKGKGTPLPTYKEAKQVLSYETLYANGSRTEXQVGIYLSSSGDAXXRDGS 120
 61 VSSSGPQCHKGKGTPLPTYKEAKQVLSYETLYANGSRTEXQVGIYLSSSGDAXXRDGS 120
 Db 121 SGKSRKRQIYGDRESFISIGKDFELLNYPESTSKLSTGCTTLAEXHVLTAACIHDG 180
 121 SGKSRKRQIYGDRESFISIGKDFELLNYPESTSKLSTGCTTLAEXHVLTAACIHDG 180
 Qy 181 KTYVGTQKLRVGFELPKFKDGGGRANDSTSAMPQEMKFWIRVKRTHVPGWIKGNAND 240
 181 KTYVGTQKLRVGFELPKFKDGGGRANDSTSAMPQEMKFWIRVKRTHVPGWIKGNAND 240
 Db 241 IGMDYDALLELKPKHKKRPMKIGVSPPAKOLPGGRHIFESGYDNDRPGNLYRFCDYKDE 300
 241 IGMDYDALLELKPKHKKRPMKIGVSPPAKOLPGGRHIFESGYDNDRPGNLYRFCDYKDE 300
 Db 301 TYDLYQQCDAQPGASGYGGYVVRMKRQQQKWERKLIGIFSPGSQHNDMNGSPQDFNVAVR 360
 301 TYDLYQQCDAQPGASGYGGYVVRMKRQQQKWERKLIGIFSPGSQHNDMNGSPQDFNVAVR 360
 Qy 361 ITPLKYAQICYWIKGNYLDREGDVFPLPSN 392
 361 ITPLKYAQICYWIKGNYLDREGDVFPLPSN 392
 Db

RESULT 2
 US-09-072-384-15
 ; Sequence 15, Application US/09072384
 ; Patent No. 6153420
 GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/072,384
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, Gary E
 REGISTRATION NUMBER: 31,648
 REFERENCE/DOCKET NUMBER: 97-16C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-442-6673
 TELEX:
 INFORMATION FOR SEQ ID NO: 15:
 LENGTH: 392 amino acids

Query Match 98.7%; Score 2085; DB 4; Length 392;
 Best Local Similarity 98.2%; Pred. No. 7.5e-217;
 Matches 385; Conservative 1; Mismatches 6; Indels 0; Gaps 0; ;
 Feature: ;
 NAME/KEY: Signal Sequence ;
 LOCATION: 1..19 ;
 OTHER INFORMATION: ;
 US-09-072-384-15

Qy 1 MAGIGLGLFLFLFLICAVGQVSPYSAWPWKTWPAVRLPVVLPQSTNLAKPDKFGAEAKLE 60
 1 MAGIGLGLFLFLFLICAVGQVSPYSAWPWKTWPAVRLPVVLPQSTNLAKPDKFGAEAKLE 60
 Db 61 VSSSGPQCHKGKGTPLPTYKEAKQVLSYETLYANGSRTEXQVGIYLSSSGDAXXRDGS 120
 61 VSSSGPQCHKGKGTPLPTYKEAKQVLSYETLYANGSRTEXQVGIYLSSSGDAXXRDGS 120
 Db 121 SGSRKRQIYGDRESFISIGKDFELLNYPESTSKLSTGCTTLAEXHVLTAACIHDG 180
 121 SGSRKRQIYGDRESFISIGKDFELLNYPESTSKLSTGCTTLAEXHVLTAACIHDG 180
 Qy 121 KTYVGTQKLRVGFELPKFKDGGGRANDSTSAMPQEMKFWIRVKRTHVPGWIKGNAND 240
 121 KTYVGTQKLRVGFELPKFKDGGGRANDSTSAMPQEMKFWIRVKRTHVPGWIKGNAND 240
 Db 181 KTYVGTQKLRVGFELPKFKDGGGRANDSTSAMPQEMKFWIRVKRTHVPGWIKGNAND 240
 181 KTYVGTQKLRVGFELPKFKDGGGRANDSTSAMPQEMKFWIRVKRTHVPGWIKGNAND 240
 Qy 241 IGMDYDALLELKPKHKKRPMKIGVSPPAKOLPGGRHIFESGYDNDRPGNLYRFCDYKDE 300
 241 IGMDYDALLELKPKHKKRPMKIGVSPPAKOLPGGRHIFESGYDNDRPGNLYRFCDYKDE 300
 Db 301 TYDLYQQCDAQPGASGYGGYVVRMKRQQQKWERKLIGIFSPGSQHNDMNGSPQDFNVAVR 360
 301 TYDLYQQCDAQPGASGYGGYVVRMKRQQQKWERKLIGIFSPGSQHNDMNGSPQDFNVAVR 360
 Qy 361 ITPLKYAQICYWIKGNYLDREGDVFPLPSN 392
 361 ITPLKYAQICYWIKGNYLDREGDVFPLPSN 392
 Db

RESULT 3
 US-09-072-384-18
 ; Sequence 18, Application US/09072384
 ; Patent No. 6153420
 GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/072,384
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, Gary E
 REGISTRATION NUMBER: 31,648
 REFERENCE/DOCKET NUMBER: 97-16C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-442-6673
 TELEX:
 INFORMATION FOR SEQ ID NO: 15:
 LENGTH: 392 amino acids

Result No.	Score	Query	Match	Length	DB ID	Description
1	2131	100.0	392	4	US-09-072-384-15	Sequence 15, Appl
2	2085	97.8	392	4	US-09-072-384-2	Sequence 2, Appl
3	2080	97.6	383	4	US-09-072-384-18	Sequence 18, Appl
4	177	8.3	222	1	US-08-090-048-1	Sequence 1, Appl
5	177	8.3	222	2	US-08-292-250-1	Sequence 1, Appl
6	177	8.3	222	2	US-07-972-661A-1	Sequence 1, Appl
7	122	5.7	256	3	US-08-906-669-89	Sequence 89, Appl
8	122	5.7	256	3	US-08-906-516-89	Sequence 89, Appl
9	122	5.7	256	4	US-08-817-795-89	Sequence 89, Appl
10	122	5.7	256	4	US-08-039-075A-89	Sequence 89, Appl
11	122	5.7	256	4	US-09-012-331-89	Sequence 89, Appl
12	122	5.7	256	4	US-09-032-215-32	Sequence 89, Appl
13	122	5.7	256	4	US-09-012-692-89	Sequence 89, Appl
14	122	5.7	256	4	US-08-906-513-89	Sequence 89, Appl
15	122	5.7	256	5	PCT-US5-1442A-89	Sequence 2, Appl
16	1117	5.5	437	1	US-08-87-037-2	Sequence 2, Appl
17	115.5	5.4	241	4	US-08-94-83-59	Sequence 59, Appl
18	114	5.3	228	4	US-08-94-483-44	Sequence 4, Appl
19	114	5.3	253	6	5223425-8	Patent No. 5223425
20	113	5.3	238	6	5223425-5	Patent No. 5223425
21	1113	5.3	250	6	5223425-4	Patent No. 5223425
22	112.5	5.3	223	1	US-08-278-091-13	Sequence 13, Appl
23	112.5	5.3	223	1	US-08-483-059-13	Sequence 13, Appl
24	112.5	5.3	223	1	US-08-47-113-13	Sequence 13, Appl
25	112.5	5.3	223	2	US-08-487-167-13	Sequence 13, Appl
26	112.5	5.3	223	2	US-08-482-816-13	Sequence 13, Appl
27	112.5	5.3	223	2	US-08-296-149-13	Sequence 13, Appl

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:15:33 ; Search time 35.78 Seconds
(without alignments)

246.542 Million cell updates/sec

Title: US-09-072-384-15
2131
Perfect score:
Sequence: 1 MAGIPGFLFLFLCAVGQ.....IKGNYLDRCGDTVFPFGSN 392

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*5: /cgn2_6/ptodata/2/1aa/PECTUS_COMB.pep:*6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
US-09-072-384-15
; Sequence 15, Application US/09072384
; Patent No. 6153420

GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
; TITLE OF INVENTION: AND MATERIALS AND METHODS FOR MAKING THEM
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072_384
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REVENUE/DOCKET NUMBER: 97-16C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; NAME/KEY: Signal Sequence
; LOCATION: 1...19
; OTHER INFORMATION:
US-09-072-384-15

Query Match 100.08; Score 2131; DB 4; Length 392;
 Best Local Similarity 100.08; Pred. No. 1..1e-27;
 Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIPGLLFLLEFLLCAGVQSPYSAFWKPTWPAVLPPVLFQSTLNALKPFGAEAKLE 60
 DB 1 MAGIPGLLFLLEFLLCAGVQSPYSAFWKPTWPAVLPPVLFQSTLNALKPFGAEAKLE 60

QY 61 VSSSGPQCHKPTPLPYYEAQYLSVETLYANGSRPETOQYIYLSSSGGQAHRDGS 120
 DB 61 VSSSGPQCHKPTPLPYYEAQYLSVETLYANGSRPETOQYIYLSSSGGQAHRDGS 120

QY 121 SGKSRRRQRTQYGDSTSIFGDLNNYPFSTSVKLSTGCTGTLVAEKHVLTAHC1HDG 180
 DB 121 SGKSRRRQRTQYGDSTSIFGDLNNYPFSTSVKLSTGCTGTLVAEKHVLTAHC1HDG 180

QY 181 KTYVKGTQKLRVGLPKFKDGGGRGANDSTSAMPQMKFQWTRKTHVPKGWIKGNAND 240
 DB 181 KTYVKGTQKLRVGLPKFKDGGGRGANDSTSAMPQMKFQWTRKTHVPKGWIKGNAND 240

QY 241 IGMDDYALLEKKPHRKEMIGVSPAKQPGGRTHFSCYDNDPQGNLYRFDQVKDE 300
 DB 241 IGMDDYALLEKKPHRKEMIGVSPAKQPGGRTHFSCYDNDPQGNLYRFDQVKDE 300

QY 301 TYDLYQCDQDAGPGASGSQGYVYRMKQROQQKWERKLIGIFSGHQMNGSPQDFNVAVR 360
 DB 301 TYDLYQCDQDAGPGASGSQGYVYRMKQROQQKWERKLIGIFSGHQMNGSPQDFNVAVR 360

QY 361 ITPLKYAQICYWKGNYLDCREGDTVFPGSN 392
 DB 361 ITPLKYAQICYWKGNYLDCREGDTVFPGSN 392

RESULT 2
 US-09-072-384-2
 ; Sequence 2, Application US/09072384
 ; Patent No. 6153420
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; TITLE OF INVENTION: AND MATERIALS AND METHODS FOR MAKING THEM
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/072, 384
 ; FILING DATE:
 ; PRIORITY APPLICATION DATA:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parker, Gary E
 ; REGISTRATION NUMBER: 31, 648
 ; REFERENCE/DOCKET NUMBER: 97-16C1
 ; TELEPHONE: 206-442-6673
 ; TELEFAX: 206-442-6678
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 392 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 FEATURE:
 NAME/KEY: Signal Sequence
 LOCATION: 1...19
 OTHER INFORMATION:
 US-09-072-384-2

Query Match 97.88; Score 2085; DB 4;
 Best Local Similarity 98.28; Pred. No. 1..5e-22;
 Matches 385; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 MAGIPGLLFLLEFLLCAGVQSPYSAFWKPTWPAVLPPVLFQSTLNALKPFGAEAKLE 60
 DB 1 MAGIPGLLFLLEFLLCAGVQSPYSAFWKPTWPAVLPPVLFQSTLNALKPFGAEAKLE 60

QY 61 VSSSGPQCHKPTPLPYYEAQYLSVETLYANGSRPETOQYIYLSSSGGQAHRDGS 120
 DB 61 VSSSGPQCHKPTPLPYYEAQYLSVETLYANGSRPETOQYIYLSSSGGQAHRDGS 120

QY 121 SGKSRRRQRTQYGDSTSIFGDLNNYPFSTSVKLSTGCTGTLVAEKHVLTAHC1HDG 180
 DB 121 SGKSRRRQRTQYGDSTSIFGDLNNYPFSTSVKLSTGCTGTLVAEKHVLTAHC1HDG 180

QY 181 KTYVKGTQKLRVGLPKFKDGGGRGANDSTSAMPQMKFQWTRKTHVPKGWIKGNAND 240
 DB 181 KTYVKGTQKLRVGLPKFKDGGGRGANDSTSAMPQMKFQWTRKTHVPKGWIKGNAND 240

QY 241 IGMDDYALLEKKPHRKEMIGVSPAKQPGGRTHFSCYDNDPQGNLYRFDQVKDE 300
 DB 241 IGMDDYALLEKKPHRKEMIGVSPAKQPGGRTHFSCYDNDPQGNLYRFDQVKDE 300

QY 301 TYDLYQCDQDAGPGASGSQGYVYRMKQROQQKWERKLIGIFSGHQMNGSPQDFNVAVR 360
 DB 301 TYDLYQCDQDAGPGASGSQGYVYRMKQROQQKWERKLIGIFSGHQMNGSPQDFNVAVR 360

QY 361 ITPLKYAQICYWKGNYLDCREGDTVFPGSN 392
 DB 361 ITPLKYAQICYWKGNYLDCREGDTVFPGSN 392

RESULT 3
 US-09-072-384-18
 ; Sequence 18, Application US/09072384
 ; Patent No. 6153420
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; TITLE OF INVENTION: SERINE PROTEASE POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/072, 384
 ; FILING DATE:
 ; PRIORITY APPLICATION DATA:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parker, Gary E
 ; REGISTRATION NUMBER: 31, 648
 ; REFERENCE/DOCKET NUMBER: 97-16C1
 ; TELEPHONE: 206-442-6673
 ; TELEFAX: 206-442-6678
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 392 amino acids

CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

Title: US-09-072-384-2	Perfect score: 2112	Sequence: 1 MGIPIGLLFFFLCAVQQ IKGNYLDCREGDTVFLPGSN 392
Scoring table: BLOSUM62	Gapop 10.0 , Gapext 0.5	
Searched: 522463 seqs, 74073290 residues		
Total number of hits satisfying chosen parameters: 522463		
Minimum DB seq length: 0		
Maximum DB seq length: 2000000000		
Post-processing: Minimum Match 0%		
Maximum Match 100%		
Listing first 45 summaries		
Database : A_Geneseq_1101;*		
1: /SIDS2/gcadata/geneseq/geneseqp/AA1980.DAT;*		
2: /SIDS2/gcadata/geneseq/geneseqp/AA1981.DAT;*		
3: /SIDS2/gcadata/geneseq/geneseqp/AA1982.DAT;*		
4: /SIDS2/gcadata/geneseq/geneseqp/AA1984.DAT;*		
5: /SIDS2/gcadata/geneseq/geneseqp/AA1985.DAT;*		
6: /SIDS2/gcadata/geneseq/geneseqp/AA1986.DAT;*		
7: /SIDS2/gcadata/geneseq/geneseqp/AA1987.DAT;*		
8: /SIDS2/gcadata/geneseq/geneseqp/AA1988.DAT;*		
9: /SIDS2/gcadata/geneseq/geneseqp/AA1989.DAT;*		
10: /SIDS2/gcadata/geneseq/geneseqp/AA1990.DAT;*		
11: /SIDS2/gcadata/geneseq/geneseqp/AA1991.DAT;*		
12: /SIDS2/gcadata/geneseq/geneseqp/AA1992.DAT;*		
13: /SIDS2/gcadata/geneseq/geneseqp/AA1993.DAT;*		
14: /SIDS2/gcadata/geneseq/geneseqp/AA1994.DAT;*		
15: /SIDS2/gcadata/geneseq/geneseqp/AA1995.DAT;*		
16: /SIDS2/gcadata/geneseq/geneseqp/AA1996.DAT;*		
17: /SIDS2/gcadata/geneseq/geneseqp/AA1997.DAT;*		
18: /SIDS2/gcadata/geneseq/geneseqp/AA1998.DAT;*		
19: /SIDS2/gcadata/geneseq/geneseqp/AA1999.DAT;*		
20: /SIDS2/gcadata/geneseq/geneseqp/AA2000.DAT;*		
21: /SIDS2/gcadata/geneseq/geneseqp/AA2001.DAT;*		
22: /SIDS2/gcadata/geneseq/geneseqp/AA2002.DAT;*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived from analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	20 AAY08657	Human transmembran
4	2044	96.8	383	20 AAY08650	W09927094 Seq ID 1
5	2044	96.8	383	20 AAY13390	Amino acid sequenc
6	2044	96.8	383	21 AAB25592	Protein encoded by
7	2044	96.8	383	21 AAB25618	Protein encoded by
8	2044	96.8	383	21 AAY88217	Human TANGO 186 pr
9	2044	96.8	383	21 AAY87270	Human signal Pepti
10	2044	96.8	383	21 AAY5327	A bone marrow secr
11	2044	96.8	383	22 AAB80258	Human PRO307 prote

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2104	99.6	392	22 AAB48972	Human Zsig13 varia
2	2098	98.7	392	AAB48973	Human Zsig13 varia
3	2044	96.8	383	2	

Scoring table:	BLOSUM62	ALIGNMENTS			
Searched:	100059 seqs, 36664827 residues	RESULT	1		
Total number of hits satisfying chosen parameters:	100059	ID	GSEP_BACLI	STANDARD;	PRT; 316 AA.
Minimum DB seq length: 0		AC	P80057;		
Maximum DB seq length: 2000000000		DT	01-MAY-1992 (Rel. 22, Created)		
Post-processing: Minimum Match 0%		DT	01-OCT-1993 (Rel. 27, Last sequence update)		
Maximum Match 100%		DT	20-AUG-2001 (Rel. 40, Last annotation update)		
Listing first 45 summaries		DE	GLUTANYL ENDOPETIDASE PRECURSOR (EC 3.4.21.19) (GLUTAMATE SPECIFIC ENDOPETIDASE) (GSE).		
Database :	SwissProt_39:*	GN	GNBLASE.		
		OS	Bacillus licheniformis		
		OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
		OC	Bacillus; Firmicutes; Bacillus/Clostridium group; Bacillus.		
		OX			
		RN	[1]		
		RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
		RC	STRAIN=ATCC 14580;		
		RX	REVIEW=;		
		RA	Medline=93054737; PubMed=1429718;		
		RA	Kakudo S., Kikuchi N., Kitadokoro K., Fujiwara T., Nakamura E., Tsuzuki H., Yoshida N.;		
		RA	Okamoto H., Shin M., Tamaki M., Teraoka H., "Purification, characterization, cloning, and expression of a glutamyl acid-specific protease from <i>Bacillus licheniformis</i> ATCC 14580".		
		RA	J. Biol. Chem. 267:23782-23788(1992).		
		RN	[2]		
		RP	SEQUENCE 95-316.		
		RX	Medline=92155199; PubMed=1346764;		
		RA	Svendsen I., Breedam K.;		
		RA	"Isolation and amino acid sequence of a glutamyl acid specific endopeptidase from <i>Bacillus licheniformis</i> .",		
		RA	Eur. J. Biochem. 204:165-171(1992).		
		RL	- FUNCTION: SPECIFIC FOR HYDROLYSIS OF PEPTIDES BONDS ON THE CARBOXYL SIDE OF ACIDIC AMINO ACID RESIDUES, WITH A STRONG PREFERENCE FOR GLU.		
		CC	- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ASP-1-XAA, GLU-1-XAA.		
		CC	-1- SUBCELLULAR LOCATION: SECRETED.		
		CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2B; ALSO KNOWN AS THE V8 FAMILY.		
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See: http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
		CC	EMBL: D10060; BAA00949.1; -.		
		DR	PTR; S23078; S23078.		
		DR	InterPro: IPR001254; Trypsin.		
		DR	Prints: PF00089; trypsin; 1.		
		DR	Prints; PRO039; V8PROTEASE.		
		DR	SMART; SM00020; tryp_SPC; 1.		

Result No.	Score	Query Match	Length	DB ID	Description
1	185	8 8	316	2 A45134	endopeptidase (EC metalloproteinase
2	130.5	6.2	313	2 A35122	coagulation factor
3	125	5.9	482	1 EXRT	pancreatic elastas
4	118	5.6	269	2 B26823	pancreatic elastas
5	114	5.4	258	1 S70439	probable pancreatic
6	114	5.4	267	4 A56615	coagulation factor
7	114	5.4	488	1 EXHU	coagulation factor
8	114	5.4	492	1 EXBO	pancreatic elastas
9	113	5.4	269	2 A26823	pancreatic elastas
10	113	5.4	271	2 A25528	trypsin (EC 3.4.21
11	112.5	5.3	259	1 TRSMG	pancreatic elastas
12	112	5.3	266	1 ELRT1	pancreatic elastas
13	112	5.3	266	1 ELPG	pancreatic elastas
14	112	5.3	522	2 T29767	hypothetical prote
15	111.5	5.3	238	1 TRW5Y	trypsin-like prote
16	111.5	5.3	686	1 A59271	Ra-reactive factor
17	111	5.3	246	1 DBHU	complement factor
18	110	5.2	269	2 C26823	pancreatic elastas
19	109	5.2	761	2 JC5759	brain-specific ser
20	107.5	5.1	405	2 T35117	probable secreted
21	106	5.0	336	2 A28566	T-cell suppressor
22	106	5.0	271	1 ELRT2	pancreatic elastas
23	104.5	4.9	273	2 BB5765	hypothetical prote
24	104.5	4.9	273	2 H64915	hypothetical prote
25	104.5	4.9	1582	2 T15308	hypothetical prote
26	102.5	4.9	2055	2 T31617	hypothetical prote
27	102	4.8	786	1 A47547	serine protease
28	101.5	4.8	583	2 A29154	complement factor
29	101.5	4.8	747	2 T151579	complement factor

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2001, 16:14:45 ; Search time 48.03 Seconds (without alignments)

621.704 Million cell updates/sec

Title: US-09-072-384-2

Perfect score: 2112

Sequence: 1 MNGIPGLLELLFELLCAVGQ IKGNYLDCREGDTVFLPGSN 392

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0⁸ Maximum Match 100⁸ Listing first 45 summaries

Database : PIR68:*

 1: PIR1:*

 2: PIR2:*

 3: PIR3:*

 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	8 8	316	2 A45134	endopeptidase (EC metalloproteinase
2	130.5	6.2	313	2 A35122	coagulation factor
3	125	5.9	482	1 EXRT	pancreatic elastas
4	118	5.6	269	2 B26823	pancreatic elastas
5	114	5.4	258	1 S70439	probable pancreatic
6	114	5.4	267	4 A56615	coagulation factor
7	114	5.4	488	1 EXHU	coagulation factor
8	114	5.4	492	1 EXBO	pancreatic elastas
9	113	5.4	269	2 A26823	pancreatic elastas
10	113	5.4	271	2 A25528	trypsin (EC 3.4.21
11	112.5	5.3	259	1 TRSMG	pancreatic elastas
12	112	5.3	266	1 ELRT1	pancreatic elastas
13	112	5.3	266	1 ELPG	pancreatic elastas
14	112	5.3	522	2 T29767	hypothetical prote
15	111.5	5.3	238	1 TRW5Y	trypsin-like prote
16	111.5	5.3	686	1 A59271	Ra-reactive factor
17	111	5.3	246	1 DBHU	complement factor
18	110	5.2	269	2 C26823	pancreatic elastas
19	109	5.2	761	2 JC5759	brain-specific ser
20	107.5	5.1	405	2 T35117	probable secreted
21	106	5.0	336	2 A28566	T-cell suppressor
22	106	5.0	271	1 ELRT2	pancreatic elastas
23	104.5	4.9	273	2 BB5765	hypothetical prote
24	104.5	4.9	273	2 H64915	hypothetical prote
25	104.5	4.9	1582	2 T15308	hypothetical prote
26	102.5	4.9	2055	2 T31617	hypothetical prote
27	102	4.8	786	1 A47547	serine protease
28	101.5	4.8	583	2 A29154	complement factor
29	101.5	4.8	747	2 T151579	complement factor

ALIGNMENTS

RESULT 1									
A45134	endopeptidase (EC 3.4.21-), glutamate-specific	-	Bacillus licheniformis						
C; Species: Bacillus licheniformis									
C; Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999									
C; Accession: A45134; S23078									
R; Kikuchi, S.; Kitadokoro, K.; Fujiwara, T.; Nakamura, E.; Okamoto, H.; S									
J. Biol. Chem. 267, 23782-23788, 1992.									
A; Title: Purification, characterization, cloning, and expression of a glutamic acid-s									
A; Reference number: A45134; MUID: 93054737									
A; Accession: A45134									
A; Status: preliminary									
A; Molecule type: DNA									
A; Residues: 1-316 <KAK>									
A; Cross-references: GB:D10060; NID:9316263; PID:BAA00949.1; PID:d1001415; PID:g21626									
A; Experimental source: ATCC 14580									
A; Note: sequence extracted from NCBI backbone (NCBIN:118784, NCBIPI:118785)									
R; Svendsen, I.; Bredam, K.									
Eur. J. Biochem. 204, 165-171, 1992									
A; Title: Isolation and amino acid sequence of a glutamic acid specific endopeptidase									
A; Reference number: S23078; MUID:92155199									
A; Accession: S23078									
A; Status: preliminary									
A; Molecule type: protein									
A; Residues: 95-116 <SYE>									
C; Keywords: hydrolase									
Query Match	8.8*	Score 185;	DB 2;	Length 316;					
Best Local Similarity	23.28;	Pred. No. 2.9e-08;							
Matches	46	Conservative	46	Mismatches 139;	Indels 70;	Gaps 16;			
Qy	63	SSCGPOQHKGTPPL--PTYKBAQYLSYETLYANGSRTEXOYGIYLSSSSDGMXNRDSG 119							
Db	77	TKEKAEGKSPAKAPYTSVGSDDRTVNN--TATAPYRAIVHISSTGSGCTGMIGP 133							
Qy	168	XHVLTAAHC1HDGKT-YVKGTQKLVRGFELPKFKDGGRGANDSTSAMPQMKFQNTRKVR 226							
Db	120	SSGSKSRK-----RQIYGDSSIFKDFLLNYPETSYKSLT--GCTGPRVAE 167							
Qy	64	SSCGPOQHKGTPPL--PTYKBAQYLSYETLYANGSRTEXOYGIYLSSSSDGMXNRDSG 117							
Db	77	TKEKAEGKSPAKAPYTSVGSDDRTVNN--TATAPYRAIVHISSTGSGCTGMIGP 133							
Qy	169	YFIPSGMRSGNTN---YDGAIESELSEPIGNTVGYGYSYTSSUVGTVNTISGPGD 230							
Qy	287	PGLVYRFCD--VKDPTDYLXQQCDAQPGASGYGYVRYMKWROQKWKERKII---GIF 340							
Db	231	TAGTQHSGDIAISETYKLOYAM-DYGGSGSPVFEQSSSRNTNCGPCLSLAVHTNGVY 289							

QY 341 SGHQWDMNGSPQDFNVAVRITPLKYAQICYW 372
 Db 290 G₁-----SSYNRQTRPKFVDFNLNW 311

RESULT 2

A35122 metalloproteinase (EC 3.4.21.6) mpr precursor, extracellular - *Bacillus subtilis*
 C;Species: *Bacillus subtilis*
 C;Date: 27-Jul-1990 #text_change 15-Oct-1999
 C;Accession: A35122; I40100; A69660
 R;Slioma, A.; Rudolph, C.F.; Rufo Jr., G.A.; Sullivan, B.J.; Theriault, K.A.; Ally, D.; J. Bacteriol. 172, 1024-1029, 1990
 A;Title: Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*.
 A;Reference number: A35122; MUID:90130256
 A;Accession: A35122
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-313 <SLO>
 A;Cross-references: GB:M10505; PIDN:q143209; PIDN:AAA22604.1; PID:9143210; GB:M29036
 R;Smith, H.; de Jong, A.; Bron, S.; Venema, G.
 Gene 70, 351-361, 1988
 A;Title: Characterization of signal-sequence-coding regions selected from the *Bacillus* s
 A;Reference number: I39994; MUID:89108019
 A;Accession: I40100
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-60, '65, '67, '67, '69, 'AQA' <RES>
 A;Cross-references: GB:M22916; PIDN:q143702
 R;Kunst, F.; Ogawa, N.; Mosser, I.; Albertini, A.M.; Alloui, G.; Arevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caprano, V.; Carter, N.M.; Chodat, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997
 A;Authors: Fouger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Galler, J.; Harwood, C.R.; Hennaut, A.; Hilbert, H.; Hollsappel, S.; Hosono, S.; Hull, M.F.; Koetter, P.; Konigstein, G.; Korch, S.; Kumano, M.; Kurita, K.; Lardinois, A.; Lardinois, Y.; M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelet, R.; Rieger, M.; Rivilta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schrofke, R.; Scroffone, F.; Sekiguchi, J.; Sejowska, A.; Seron, A.; Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Toononi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshihikawa, H.F.; Zumstein, E.; Yoshihikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A;Reference number: A653580; MUID:98044033
 A;Accession: A69660
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-313 <KUN>
 A;Cross-references: GB:AL009126; PIDN:q2632457; PIDN:CAIB12018.1; PID:e1182176;
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: mpr
 C;Keywords: hydrolase

Query Match 6.2%; Score 130.5; DB 2; Length 313;
 Best Local Similarity 22.6%; Pred. No. 0.0013; Indels 121; Gaps 19;

Matches 87; Conservative 38; Mismatches 139; Del 61

28 WKPPTWPAVLPPVLPQSTNLAKPDKFGEAEAKLEVS----SSCGPQCHKGKTPLYKEAK 82
 Db 7 FPKQWFAFLTVLCLALA----AAYSFGVPAKAAENPQTISVNTGKEA-DATKNQTSKADQ 61

QY 83 QYQSYE-----TLYANGSRPEXQVGYIYLSGSGSKSRRQIYGDS 135
 Db 62 VSAPYEGTNGKTSKSLY--GGQPFLEKNIQTLQS-----SIIGTDE 100

QY 136 RFSIFGKDFELNYPFESTPSVYKST-----GCTGTLVAEXHVLTAACHT-----178
 Db 101 RTRR-----SSTTSVPPYRAVQLSLKYPNTSSTVGTGFIVNPNTVVTAGCHYYSDHHWAS 157
 QY 179 -----DGKTTVKGQKQLRNGFLKPKFKDGGRGANDSTSAMPQMKFQWIRVKRTHVP 230

Db 158 TITAAPGNGSSSPYGY 195
 QY 231 KGWTIKGN---ANDGMDYDYLLELKKHKRKEMKIGSPPAKOLPGRTHFSGYDNRP 287
 Db 196 YGA1KLNSPGNTVGVW YGYRITNSSL-----VGUSSSTGFPCKRTGTMWSDFKRP 247

QY 288 GNLYVRFCDVKEDETYLQQCDAQPGASGYVVRMKRQQQKWERKLIGIFSGHQWVD 347
 Db 248 IR-----SAETYKLY-TDITYGCGSGSPVY----RNYSDTGOTAIHT-----287
 QY 348 MNGSPQDINAVVIRTPLYAQICYW 372
 Db 288 -NGG-SSYNLGTRVNDVNNIQYW 310

RESULT 3

EXRT coagulation factor Xa (EC 3.4.21.6) Precursor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Accession: S41905; JC4670; PS0191; PS0190; I62745
 R;Stanton, C.; Ross, P.; Huisson, S.; Wallin, R.
 Thromb. Res. 80, 63-73, 1995
 A;Title: Evidence for competition between vitamin K-dependent clotting factors for in
 A;Reference number: MUID:96093366
 A;Accession: S4015
 A;Molecule type: mRNA
 A;Residues: 1-482 <STAA2>
 A;Cross-references: EMBL:X79807; PIDN:CAA56202.1; PID:9506601
 A;Accession: S4015; JC4670; PS0191; PS0190; I62745
 A;Note: submitted to the EMBL Data Library, June 1994
 A;Title: Processing and expression of rat and human clotting factor X-encoding cDNAs.
 Gene 169, 269-273, 1996
 A;Reference number: MUID:96194815
 A;Molecule type: mRNA
 A;Residues: 1-482 <STAA2>
 A;Cross-references: EMBL:X79807; PIDN:9506600; PIDN:CAA56202.1; PID:9506601
 A;Experimental source: Cos-1 cell
 R;Enjyoji, K.; Mizraaki, K.; Kato, H.
 J. Biochem. 109, 890-898, 1991
 A;Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat
 A;Reference number: PS0190; MUID:92041742
 A;Accession: PS0191
 A;Molecule type: protein
 A;Residues: 41-58 'X', 60-65 <ENJ1>
 A;Accession: PS0190
 A;Molecule type: protein
 A;Residues: 183-186 'X', 188-207 <ENJ2>
 A;Cross-references: GB:D21215; PID:94515309; PIDN:BAA04756.1; PID:9455396
 C;Function:
 A;Pathway: blood coagulation
 C;Superfamily: coagulation factor X; EGFR homology; Gla domain homology; trypsin homolog
 C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-40/Domain: propeptide #status predicted <PRO>
 F;25-84/Domain: Gla domain homology <GLA>
 F;41-179/Domain: coagulation factor X light chain #status predicted <LCX>
 F;50-123/Domain: EGFR homology <EG1>
 F;129-167/Domain: EGFR homology <EG2>
 F;183-482/Domain: coagulation factor X heavy chain #status predicted <HCH>
 F;183-231/Domain: activation peptide #status predicted <AP1>